[P]UBLICATIONS DE L'INSTITUT DE STATISTIQUE DE L'UNIVERSITÉ DE PARIS

MÉMOIRES ET CONFÉRENCES SUR LE CALCUL DES PROBABILITÉS, LA STATISTIQUE THÉORIQUE ET APPLIQUÉE, L'ÉCONOMÉTRIE

VOL. III - FASCICULE 3 - 1954

PARIS

11, Rue Pierre Curie

Toute la correspondance relative aux publications
doit être envoyée à l'adresse :

INSTITUT DE STATISTIQUE DE L'UNIVERSITE DE PARIS

INSTITUT HENRI POINCARÉ - 11, Rue Pierre Curie - PARIS (Ve)

Les manuscrits doivent être envoyés à *M. Daniel DUGUÉ,*
à l'adresse précédente.

Abonnements :	Pour la France 1.200 francs français
	Pour l'Etranger 1.500 francs
Vente au numéro :	*(fascicule de 50 pages environ)*
	Pour la France 350 francs français
	Pour l'Etranger 400 francs

UBLICATIONS DE L'INSTITUT DE STATISTIQUE
DE L'UNIVERSITÉ DE PARIS

MÉMOIRES ET CONFÉRENCES SUR LE CALCUL DES PROBABILITÉS,
LA STATISTIQUE THÉORIQUE ET APPLIQUÉE, L'ÉCONOMÉTRIE

VOL. III - FASCICULE 3 - 1954

PARIS
11, Rue Pierre Curie

PROBLÈMES DE L'ANALYSE SPECTRALE DES SÉRIES TEMPORELLES STATIONNAIRES

par

M. S. BARTLETT

1. INTRODUCTION : QUELQUES REMARQUES

L'analyse statistique des séries temporelles, et des processus stochastiques en général, ne diffère pas en principe de l'analyse des autres types de calcul statistique. Toutefois, il est bien connu que la déduction statistique doit dépendre en premier lieu des spécifications théoriques appropriées; aussi n'est-il pas surprenant que de récents développements dans la spécification théorique des processus stochastiques, et en particulier des séries temporelles stationnaires, puissent avoir une répercussion considérable sur les méthodes de l'analyse statistique. Il est également bien apparent que, pour les processus stochastiques, ce n'est que ces dernières années qu'on a commencé à prêter une sérieuse attention aux rapports qui existent entre la théorie et la méthode. Il y a une publication fondamentale, bien que très abstraite, qui est de Grenander (1950). Le travail des mathématiciens anglais, avec la note de Yule qui a devancé les autres (1927) sur les séries temporelles, tend à être d'un caractère en quelque sorte plus utilitaire. Suivant la voie tracée par Yule, il y a eu une tendance à passer de l'analyse classique du type de "périodogramme" associée au spectre des séries, à une analyse d'autocorrélation et d'auto-régression associée au "corrélogramme" des séries (voir, par exemple, M.G. Kendall, 1946; Bartlett, 1946). Cela est en accord raisonnable avec l'expérience: bien des séries temporelles appartiennent au type qui présente un spectre continu dans le sens du théorème de Wiener-Khintchine (à voir dans Levy, 1948) et du théorème analogue sur les séries discrètes énoncé par Wold; et, de plus, ces mêmes séries ont une représentation de base autorégressive (Wold 1938) qui peut être finie avec quelques termes seulement, au moins en faisant une approximation. Pour ces modèles finis autorégressifs, et (pour les valeurs discrètes du temps) de la forme:

$$X_t + a_1 X_{t-1} + \dots\dots a_k X_{t-k} = Y_t$$

des solutions asymptotiques aux problèmes d'estimation et de la qualité d'ajustement existent pour les résidus normaux Y_t (en supposant qu'on a un écart σ^2 constant et une moyenne nulle), ou, plus généralement, pour les résidus indépendants. Ainsi Mann et Wold (1943) ont donné une solution très satisfaisante du problème d'estimation, aussi bien pour les séries simples que

pour les séries multiples; et des tests de "goodness of fit" ont été exposés par divers auteurs (Quenouille, 1947; Wold, 1949; Bartlett et Diananda, 1950 ; Whittle 1951, 1952, Walker, 1952; et pour les séries multiples, Bartlett et Rajalakshman, 1953; Whittle, 1953; voir pour une revue générale, Bartlett, 1954). Somme toute, la tendance à tout rapporter aux méthodes de corrélation a été jusqu'à un certain point une réaction contre l'insuccès de la technique classique du périodogramme pour ces séries, et contre la découverte, faite parfois empiriquement, que tout essai direct d'estimation de la fonction spectrale engendre des fluctuations qui augmentent, au lieu de diminuer, avec la longueur de la série. Avant de dire le mot final sur les valeurs relatives des méthodes de corrélation ou de périodogramme, il est prudent de revoir cette dernière avec plus de soin. On peut naturellement remarquer que du théorème de Wiener-Khintchine l'analyse de corrélation et l'analyse harmonique sont maintenant bien connues comme un double aspect de la même analyse théorique (on peut montrer qu'il en est de même pour l'analyse arithmétique de la série qui se réalise) et qu'alors la méthode employée est en fait peu importante, supposant toujours une interprétation juste. Il serait prématuré de chercher à porter dès maintenant un jugement définitif, mais il est peu probable que ce soit si simple; le problème de dégager les caractéristiques de données statistiques au moyen de quelques constantes, dans le cas de séries temporelles, se résout au mieux tantôt par les corrélations, tantôt par les périodogrammes. Ainsi, les constantes d'un modèle fini autorégressif seront mieux déterminées par les corrélations; et les termes discrets harmoniques dans une série harmonique classique par le périodogramme. Je me propose ici de discuter de certains problèmes restés en suspens, associés à des périodogrammes ou à des analyses spectrales et j'indiquerai des solutions qui sont plutôt possibles que finales.

2. PÉRIODOGRAMMES ET SPECTRES CONTINUS

Pour une série temporelle réelle strictement stationnaire en temps discret(1) X_t , le périodogramme classique est obtenu en calculant (soit pour l'intégrale p) la quantité:

$$G_p \equiv A_p + i B_p = \sqrt{\frac{2}{n}} \sum_{r=1}^{n} X_r \, e^{2\pi i pr/n}$$

(on calcule A_p et B_p séparément), et l'intensité $I_p = G_p G_p^* = A_p^2 + B_p^2$ Le facteur $\sqrt{(2/n)}$ a été introduit pour que la moyenne $E\{I_p\}$ soit égale à $2\sigma^2$ pour une série à termes complètement indépendants. Nous supposons pour simplifier, que $E\{X_t\} = 0$, bien que cela ne soit pas essentiel pour $p \neq 0$. Si maintenant nous définissons la fonction de covariance

$$C_s = \frac{1}{n-s} \sum_{r=1}^{n-s} X_r X_{r+s} , \qquad (s > 0 , C_{-s} = C_s)$$

(1) L'analyse d'une série temporelle en temps continu est naturellement semblable en principe: pour une discussion plus serrée de ce sujet, voir Davenport, Johnson and Middleton (1952)

un calcul assez simple nous montre que:

$$I_p = 2 \sum_{s=-n+1}^{n-1} \left(1 - \frac{|s|}{n}\right) C_s \cos(s\omega_p), \qquad \left(\omega_p = \frac{2\pi p}{n}\right)$$

C'est la relation entre I_p et le corrélogramme C_s/C_o dont on a déjà parlé; c'est cependant la moyenne de cette relation, soit

$$E\{I_p\} = 2\sigma_X^2 \sum_{s=-n+1}^{n-1} \left(1 - \frac{|s|}{n}\right) \rho_s \cos(s\omega_p)$$

qui correspond pour $n \to \infty$ à la relation de Wold entre ρ_s et la fonction spectrale. Quand cette dernière est absolument continue avec, pour la densité $f(\omega)$, $0 \leqslant \omega \leqslant \pi$, cette relation devient

$$2\pi\sigma_X^2 \; f(\omega_p) = 2\sigma_X^2 \sum_{s=-\infty}^{\infty} \rho_s \cos(s\omega_p)$$

Considérons maintenant les fluctuations de I_p.

On sait que, pour une suite normale indépendante Y_t, ces fluctuations sont représentées par la loi de probabilité:

$$P\{I_p \geqslant Z\} = \exp\left(-1/2 \frac{Z}{\sigma^2}\right)$$

et, en particulier, $\mathrm{var}(I_p) = 4\sigma^4 = [E\{I_p\}]^2$. Même si Y_t n'est pas normale, nous avons encore

$$E\{H_p H_q^*\} = \frac{2\sigma_Y^2}{n} \; \frac{\left(1 - e^{in(\omega_1 \pm \omega_2)}\right) e^{i(\omega_1 \pm \omega_2)}}{1 - e^{i(\omega_1 \pm \omega_2)}}$$

où $\omega_1 = 2\pi p/n$, $\omega_2 = 2\pi q/n$, le signe $\pm$ correspond à H_q, H_q^*, et $H_p = \sqrt{(2/n)} \sum Y_r \, e^{2\pi i p \, r/n}$. Celle-ci disparait avec p et q entiers, ce qui montre que H_p n'est en corrélation ni avec H_q ni avec H_q^*, ou ce qui revient au même, que les parties réelles et imaginaires de H_p et H_q ne sont pas en corrélation. $E\{H_p^2\}$ est aussi nulle, donc les parties réelles et imaginaires de H_p ne sont pas en corrélation et apportent le même carré moyen dans $E\{H_p \; H_p^*\}$.

De plus, pour les quantités du 2° ordre $J_p = H_p H_p^*$ on trouve immédiatement que $E\{J_p \; J_q\} - 4\sigma_Y^4$

$$= \frac{4K_4}{n} + \frac{4\sigma_Y^4}{n^2} \left[\frac{1 - \cos n(\omega_1 + \omega_2)}{1 - \cos(\omega_1 + \omega_2)} + \frac{1 - \cos n(\omega_1 - \omega_2)}{1 - \cos(\omega_1 - \omega_2)} \right]$$

où $K_4 = E\{Y_t^4\} - 3\sigma_Y^4$

Ceci conduit aux conclusions suivantes:

(i) $\mathrm{var}\{J_p\} \sim 4\sigma_Y^4$ ce résultat étant toujours vrai si Y_t n'est pas normale

(ii) $\mathrm{covar}\{J_p \; J_q\}$, pour $n|\omega_1 - \omega_2| \gg 1$, est:

a) $O(1/n^2)$ pour Y_t normale (exactement zéro si p, q sont entiers)

b) $O(1/n)$ si $K_4 \neq 0$.

Supposons maintenant que X_t soit ce que j'ai appelé (Bartlett 1946) un processus linéaire

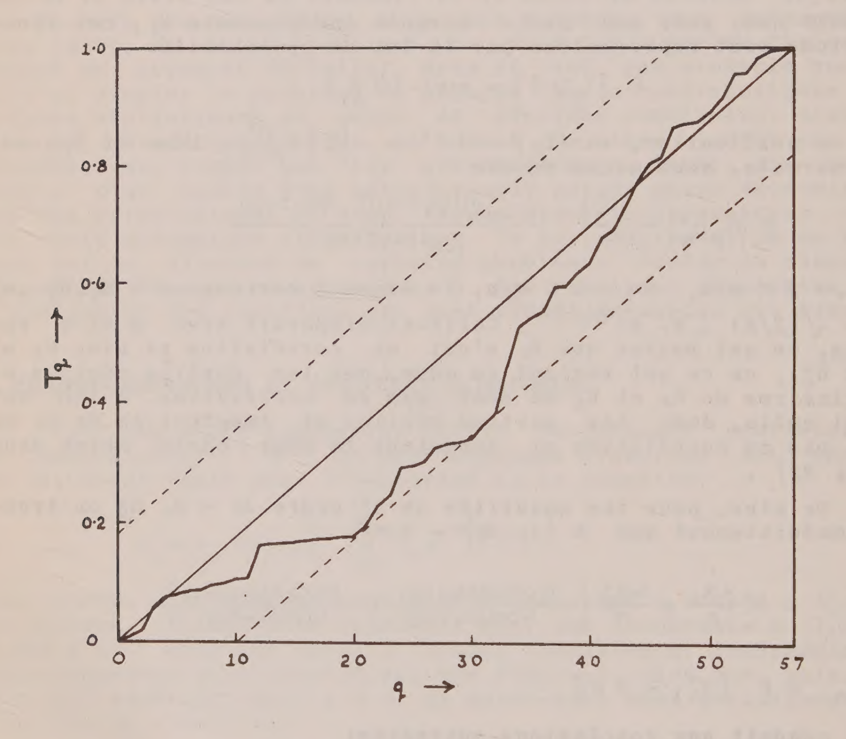
$$X_t = \sum_{u=-\infty}^{\infty} b_u Y_{t-u} \qquad (b_u = 0 \text{ pour } u < 0),$$

Cette représentation et la représentation autorégressive pour Y_t orthogonale (sans corrélation) sont équivalentes, mais l'hypothèse supplémentaire c'est que Y_t est une suite normale ou au moins indépendante. Alors

$$G_p = \sqrt{\frac{2}{n}} \sum_{u=-\infty}^{\infty} \sum_{r=1}^{n} e^{ir\omega} b_u Y_{r-u}, \quad (\omega = 2\pi p/n)$$

$$= \beta^*(\omega) H_p \{1 + 0(1\sqrt{n})\}$$

où $\beta(\omega) = \sum_{u=-\infty}^{\infty} b_u e^{-iu\omega}$



On suppose que $b_u \to 0$, exponentiellement, lorsque u croit, la différence entre G_p et $\beta^*(\omega) H_p$ n'étant dûe qu'à "l'effet de l'extrémité" (end effect). En fait on montre que ceci affecte la distribution de G_p (ou I_p) jusqu'à seulement $O(1/n)$ (cf. Sargan, 1953), mais le résultat plus faible pour le résidu découle plus directement, et il est suffisant pour notre étude.

Le résultat asymptotique

$$G_p \sim \beta^*(\omega) H_p$$

a une importance considérable. Il nous donne tout de suite la relation asymptotique

$$I_p \sim \beta(\omega)\,\beta^*(\omega)\,J_p \qquad (\beta(\omega)) \text{ supposé non nul})$$

Si nous prenons la moyenne de cette relation, nous avons pour $n \to \infty$, le résultat théorique

$$2\pi\sigma_X^2\, F(\omega) = 2\sigma_Y^2\, \beta(\omega)\, \beta^*(\omega)$$

mais ici l'on a besoin de la relation sous la forme stochastique Cela nous donne la relation asymptotique entre le périodogramme d'un processus linéaire et le périodogramme d'une suite indépendante avec un spectre uniforme et nous montre que la distribution de I_p est encore asymptotiquement valable (au moins pour Y_t normale) à condition de remplacer $2\sigma^2$ par $\lim_{n\to\infty} E\{I_p\}$ .Ce résultat a été indiqué d'abord par Slutsky (1927) (bien que pas très clairement).

Il est évident d'après les résultats énoncés ci-dessus, que I_p ne saurait fournir une estimation pour $2\pi\sigma_X^2\, f(\omega)$ bien que C_s/C_0 soit une estimation de ρ_s. On a été amené ainsi à suggérer certaines modifications pour avoir des estimations"consistantes" (cf. Bartlett, 1950; cf. Tukey, 1949), par exemple,

(i) $$I'(\omega) = \frac{1}{2\varepsilon}\int_{\omega-\varepsilon}^{\omega+\varepsilon} I(\omega)\, d\omega,$$ (P.J. Daniell)

(ii) $$\bar{I}(\omega) = 2\sum_{s=-n'+1}^{n'-1}\left(1-\frac{|s|}{n'}\, C_s \cos(s\omega)\right)$$ (Bartlett)

où $n = n'm$. Cette dernière formule présente la propriété utile de permettre le calcul à partir des $n' - 1$ premières autocorrélations de la série réalisée et ne nécessite pas une évaluation complète de I_p.

Grenander (1951) a ensuite considéré la formule plus générale, avec une fonction pondérale arbitraire $\alpha(\omega)$,

(iii) $$I''(\omega) = \int I(\omega)\,\alpha(\omega)\, d(\omega)$$

qui comprend (i) et (ii) comme cas particuliers. On ne peut faire un bilan final de l'efficacité pratique des différentes méthodes.

(Un étudiant de Manchester,J. Medhi, dans un travail en cours, a entrepris la recherche de la fonction optimum $\alpha(\omega)$ de $I''(\omega)$ dans des conditions déterminées)

3. BANDE ASYMPTOTIQUE DE CONFIANCE DE LA FONCTION SPECTRALE

Grenander a aussi attaqué cette intéressante question: la construction de la bande asymptotique de confiance de la fonc-

tion spectrale entière, et, (en collaboration avec Rosenblatt, 1952,1953), il a proposé une solution dans certaines conditions. J'ai l'impression qu'on peut avoir des sòlutions plus satisfaisantes si l'on accorde une considération adéquate aux résultats précédents sur les réalisations et je me propose donc d'étudier ce problème

La transformation linéaire de X en $A_p B_p$ pour $p=1,\ldots\frac{1}{2}(n-1)$, et A_0 pour n impair; où $p=1,\ldots\frac{1}{2}(n-2)$, et $A_0\ A_{1/2n}$, pour n pair, est orthogonale (en admettant que le facteur 2 dans $\sqrt{2}/n$ est supprimé pour p=0 ou 1/2 n). Nous remarquons que de plus:

$$\sum_{r=1}^{n} X_r^2 = \sum_p \left(A_p^2 + B_p^2\right)$$

$$\sim \sum_{p=0}^{\frac{1}{2}n} I_p \qquad \left(\frac{1}{2}(n-1) \text{ au lieu de } \frac{1}{2}n \text{ si n est impair}\right)$$

maintenant pour une suite normale indépendante X_r, il s'ensuit que la somme cumulative

$$S_q = \sum_{p=0}^{q} I_p$$

représente une promenade au hasard. Si nous imposons la condition que:

$$\sum_{r=1}^{n} X_r^2 \sim \sum_{p=0}^{\frac{1}{2}n} I_p = S_{\frac{1}{2}n}$$

soit donné, il ressort de la théorie des distributions laplaciennes que la distribution de S est équivalente à celle de

$$T_q = S_q / S_{\frac{1}{2}n} \qquad (\text{indépendamment de } S_{\frac{1}{2}n})$$

T_q variant de 0 à 1 lorsque q varie de 0 à $\frac{1}{2}$ n. Les T_q représentent une promenade au hasard conditionnelle qui peut être testée contre son espérance mathématique 2q/n selon la théorie de distribution des fonctions de répartition énoncée d'abord par Kolmogoroff (pour son interprétation selon la théorie de marche au hasard, voir par exemple Bartlett, 1949). Ainsi l'écart de T_q à $2q/n$ n'excède jamais $\pm\lambda/\sqrt{(\frac{1}{2}n)}$ avec la probabilité

$$\sum_{s=-\infty}^{\infty} (-1)^s e^{-2\lambda^2 s^2}$$

Ce test est asymptotiquement valable pour X_r indépendante et normale. Mais là-dessus nous avons édifié un critère qui est un rapport et donc une fonction des autocorrélations réalisées $R_s = C_s/C_0$; il ne sera pas donc sensible à la non normalité (cf. Bartlett, 1946); en particulier la variance asymptotique de T_q

sera indépendante de K_4, le cumulant du quatrième ordre de X_r. En fait on peut vérifier par nos résultats précédents sur les réalisations que

$$\mathrm{var}\left\{T_q\right\} \sim 4q\left(1-2q/n\right)/n^2$$

indépendamment de K_4 , et en conformité avec une marche au hasard restreinte à

$$T_{1/2n}=1$$

Pour adapter ce test au processus linéaire X_r, nous remarquons que:

$$\sum_{p=o}^{q} I_p / \left[\beta(\omega)\beta^*(\omega)\right] = \sum_{p=o}^{q} J_p\left(1+0\left(1/n\right)\right) \qquad \left(\omega=\frac{2\pi p}{n}\right)$$

(l'ordre de grandeur du résidu reste le même que précédemment puisque $J_p > 0$ et que son espérance mathématique est la même pour des valeurs différentes de ω). Par suite, finalement, comme $\beta(\omega)$ $\beta^*(\omega)$ est proportionnel à f (ω), nous pouvons prendre

$$T_q \sim \sum_{p=o}^{q} \frac{I_p}{f(\omega)} \Big/ \sum_{p=o}^{1/2n} \frac{I_p}{f(\omega)}$$

Cette fonction est immédiatement utilisable pour tester la qualité de l'ajustement d'une fonction spectrale quelconque donnée a priori. Elle est en principe également utilisable pour construire une bande de confiance pour un spectre inconnu; mais puisque cette fonction dépend de f(ω) d'une façon compliquée au point de vue du calcul arithmétique, elle est au moins commode pour cette construction. Pour éviter la pondération de chaque I_n on peut réciproquement considérer la fonction non pondérée

$$U_q = \sum_{p=o}^{q} I_p \Big| \sum_{p=o}^{1/2n} I_p$$

Malheureusement les propriétés de la répartition de U_q ne sont pas simples, car U_q n'est plus (même pour les processus laplaciens) équivalente à une quantité associée à une promenade au hasard (à moins que son dénominateur soit considéré comme fixe). Sa structure suppose, cependant, que sa distribution est insensible à la non-normalité; sa variance asymptotique, en particulier, est obtenue comme

$$\frac{4\pi^2}{n^2}\left[\left(1-F(\omega_q)\right)^2 \sum_{p=o}^{q} f^2(\omega) + F^2(\omega_q) \sum_{p=q+1}^{1/2n} f^2(\omega)\right]$$

où nous avons écrit

$$\sum_{p=o}^{q} f(\omega) \Big/ \sum_{p=o}^{1/2n} f(\omega) \sim F(\omega_q)$$

Si, ensuite, nous réduisons U_q en divisant son écart de F (ω_q) par

$$2\pi\sqrt{\left[\frac{1}{2n}\sum_{p=o}^{1/2n} f^2(\omega)\right]} \sim \sqrt{\left[2\pi\int_o^{\pi} f^2(\omega)\, d(\omega)\right]} = \sqrt{v}\ ,$$

nous suggérons que la quantité $\left[U_q - F(\omega_q)\right]/\sqrt{v}$, pourrait (comme approximation temporaire) en attendant de nouvelles recherches sur sa répartition[1] être testée de la même manière que $T_q - 2q/n$

Un tel critère alternatif pour T_q ne dépend toujours que de la fonction spectrale $f(\omega)$, mais il est peut être commode d'introduire une estimation consistante fonction de l'échantillon,et cela ne changera pas le test de U_q lorsque n est grand. Une telle estimation est, par exemple, obtenue par (1)

$$\sum_{p=0}^{1/2n} I^2(\omega) / \left[\sum_{p=0}^{1/2n} I(\omega) \right]^2$$

à condition de se rappeler que la valeur moyenne de $I^2(\omega)$ est asymptotiquement le double du carré de $E\{I(\omega)\}$, de sorte que la quantité ci-dessus a la valeur asymptotique

$$\frac{n}{2\pi} \int_0^{\pi} 2 f^2(\omega)\, d\omega \Big/ \left(\frac{n}{2\pi}\right)^2 = 4v/n$$

La fonction proposée et étudiée par Grenander et Rosenblatt est

$$V_q = \int_0^{\omega_q} I(\omega)\, d\omega \sim \frac{2\pi}{n} \sum_{p=0}^{q} L(\omega) ;$$

mais elle me paraît avoir le désavantage, sa dépendance de K_4 mise à part, d'être une estimation de la "fonction spectrale" $\sigma_X^2 F(\omega_q)$ correspondant à l'autocovariance $\sigma_X^2 \rho_s$; alors qu'elle est généralement moins intéressante que la fonction spectrale définie ici, $F(\omega_q)$, correspondant directement au corrélogramme ρ_s. Théoriquement, T_q semble supérieur à U_q ou à V_q. Une démonstration rigoureuse des propriétés asymptotiques de V_q a été donnée par Grenander et Rosenblatt ; une démonstration analogue doit être bien plus simple pour T_q.

Il peut parfois être plus utile d'avoir des tests qui n'utilisent pas les "sommes cumulatives", mais utilisent toujours les propriétés asymptotiques de I_p. Un désavantage des tests précédents semble être que si I_p s'échelonne sur un spectre uniforme comme dans T_q, alors ω a le même poids dans tout son domaine de variation; bien que (pour employer la terminologie de la physique) "l'énergie" du spectre peut être largement concentrée dans un domaine particulier. D'autre part, si I_p ne s'échelonne pas,

---

(1) La vraie variance asymptotique de U_q peut, pour toute fonction particulière $f(\omega)$ être comparée, s'il en est besoin, avec la variance entraînée par l'approximation (une telle comparaison pour l'exemple numérique du § 5 a montré que l'approximation était dans ce cas tout à fait acceptable). La validité de l'emploi de U_q au lieu de T_q peut, cependant, être beaucoup accrue en modifiant $f(\omega)$ pour la rapprocher de l'uniformité au moyen d'un changement d'échelle en divisant par une densité provisoire $f_0(\omega)$. Ce procédé donne une statistique effectivement équivalente à T_q, mais pour laquelle une bande de confiance peut être immédiatement obtenue.

Cette estimation ne serait pas très juste, si la densité du spectre présentait une pointe importante pour quelque ω car son erreur d'échantillonnage serait alors grande sauf pour n très grand.

comme c'est le cas avec U_q et V_q , la moyenne supérieure de I_p dans les domaines plus importants de ω tendra à réduire la vitesse d'approche vers ces propriétés asymptotiques.

On a les deux alternatives:

i) Tester le plus grand I_p après avoir pondéré à l'uniformité au moyen d'un $f(\omega)$ hypothétique; cependant, ce test là semblera être plus sensible à un élément spectral anormal discret, et naturellement il est employé dans le cas "classique" de $f(\omega)$ déjà uniforme (Fisher, 1929; cf. voir aussi le § 4 de cette note).

ii) Tester l'homogénéité de I_p au moyen d'un χ^2 pour l'homogénéité des variances (Bartlett, 1937; voir aussi Bartlett, 1950, Whittle 1952 a, dans le présent contexte). Si les I_p sont employés dans sommation, chacun d'eux est asymptotiquement équivalent à une estimation de variance avec deux degrés de liberté. Après la pondération, ces estimations doivent être homogènes. Ce test est très souple, car avant de faire le test on peut faire un groupement de I_p pour un domaine de ω peu important, afin de le rendre plus sensible et tout aussi valable pour le domaine plus important de ω . Une conséquence intéressante des propriétés d'échantillonnage de I_p déjà indiquées c'est que ce test doit être insensible à l'absence de normalité dans le présent contexte (1) .

4. SPECTRES MIXTES

Jusqu'à présent nous avons considéré des séries temporelles ayant des spectres continus, - en contraste avec des séries harmoniques classiques avec quelques fréquences discrètes, à part peut-être un élément "d'erreur" au spectre uniforme. Dans de nombreuses applications il est raisonnable de supposer que l'on sait *a priori* à laquelle de ces deux classes principales appartient la fonction spectrale. Il faut remarquer que tout élément harmonique discret à fréquences connues peut être éliminé par la méthode classique des moindres carrés; une variation annuelle ou saisonnière tomberait dans cette catégorie. La présence de tout spectre continu complique l'erreur d'estimation d'amplitude et de phase, mais pas plus que cela ne se produit dans l'estimation d'une moyenne générale non nulle.

En somme, il reste encore à discuter deux problèmes qui ne se présentent heureusement pas très souvent. Dans le premier nous pouvons avoir, ou bien un spectre discret à "bruit" uniforme ou bien un spectre continu; dans le second nous pouvons avoir les deux simultanément. Pour le premier la difficulté réside en ce que pour une longueur finie de séries, l'intensité du périodogramme près de la fréquence d'un élément discret peut être confondu avec le maximum d'un élément spectral continu et vice-versa. C'est un problème de discrimination où, si nous connaissons la distribution des résidus, le critère convenable est le

---

(1) On suppose remplie la condition que le nombre de degrés de liberté de χ^2 est restreint, disons, à un nombre fixé n_0 en groupant les I_p . (L'emploi du nombre maximum $\frac{1}{2}n$ a été en effet considéré par Wittle, 1952 a; mais, même dans le cas normal, sa discussion m'apparaît avoir besoin de développements supplémentaires).

quotient de vraisemblance p_0/p_1 ; où p_0 et p_1 sont les vraisemblances des deux hypothèses respectives. La difficulté réside en ce que ni l'une ni l'autre hypothèse alternative n'est exactement connue, mais renferme plusieurs constantes non déterminées. On peut toutefois tenir compte approximativement de l'effet produit par l'utilisation des estimations de ces constantes. Ainsi si nous voulons comparer les deux hypothèses H_0, H_1 le processus normal consiste à voir dans:

H_0: un élément spectral uniforme continu avec un seul élément harmonique discret surajouté

H_1: une série autorégressive du second ordre, nous avons alors:

$$(\log p_0)_{max} = C - \frac{1}{2} n \log \sum_0 ,$$

où $\sum_0$ désigne la somme des carrés des résidus après avoir ajusté l'élément discret harmonique (on suppose que la variance résiduelle σ^2 est inconnue). De même

$$(\log p_0)_{max} = C - \frac{1}{2} n \log \sum_1 ,$$

où $\sum_1$ est la somme des carrés des résidus après avoir ajusté la relation autorégressive. Comme $\sum_0$ perd deux degrés de liberté quand on ajuste un élément discret de fréquence connue (l'erreur supplémentaire provenant de l'estimation de fréquence peut être négligée) et, qu'il en est de même pour $\sum_1$, nous pouvons provisoirement admettre l'effet d'estimation en écrivant

$$\log p_0 - \log p_1 = - \frac{1}{2} (n-2) \log \left(\sum_0 / \sum_1 \right)$$

(on suppose que, à l'origine, le nombre d'observations dans la fonction de vraisemblance est n dans chaque cas; souvent, par commodité, les premières K observations ne sont pas employées dans l'ajustement d'une série autorégressive d'ordre K; alors il est utile de faire un ajustement de plus pour rendre p_0 et p_1 aussi comparables que possible).

La manière la plus simple de tester le quotient précédent consiste à employer l'adaptation suggérée par G.A. Barnard de la théorie de Abraham Wald de tests séquentiels. Nous regardons notre échantillon de n observations comme le premier d'une suite hypothétique d'échantillons indépendants et adoptons la règle suivante;

a) si $p_0/p_1 > (1-\varepsilon)/\varepsilon$ on adoptera H_0

b) si $p_0/p_1 < \varepsilon/(1-\varepsilon)$ on adoptera H_1

c) si $(1-\varepsilon)/\varepsilon > p_0/p_1 > \varepsilon/(1-\varepsilon)$ on ne décidera rien sans de nouvelles données.

Ainsi le risque d'erreur d'accepter H_0, si H_1 est vrai ou vice-versa, n'est au moins pas supérieur à ε (1)

---

(1) Cette règle obligera à de trop prudentes évaluations des risques si (comme nous l'espérons) la probabilité d'une décision avec un seul échantillon disponible n'est pas petite. Quelques exemples explicites ont été étudiés par J. Medhi.

Le dernier problème que nous avons cité, et où l'on a à la fois le spectre discret et le spectre continu est la question la plus difficile à résoudre. Wittle (1952) suggère comme méthode possible une standardisation approximative du périodogramme d'échantillon I_p par $f(\omega)$ (comme dans notre fonction précédente T_q).

Nous avons vu que ceci réduit asymptotiquement le périodogramme d'une série à spectre continu à un périodogramme d'une série à spectre uniforme, de sorte qu'on peut employer la méthode "classique" de détection des éléments discrets sur des bruits uniformes du fond. La difficulté réside naturellement dans ce fait que nous ne connaissons pas $f(\omega)$; mais Whittle suggère qu'elle peut être estimée par les procédés de régularisation que nous avons déjà étudiés dans ces conférences. Comme cette estimation peut être influencée à son tour par des éléments discrets non encore éliminés, une ou plusieurs itérations de ce procédé peuvent être nécessaires. Même alors, il n'est pas très clair jusqu'à quel point cette itération converge pour des séries finies. Comme je l'ai déjà souligné, de nouvelles recherches sur tous ces problèmes que nous venons de discuter sont encore nécessaires avant qu'on puisse enfin imposer leur valeur pratique.

5. APPLICATIONS NUMÉRIQUES

Nous allons appliquer quelques uns des tests des paragraphes 3 et 4 à des records de chasse au lynx canadien (cf. Moran 1937); ces records montrent un "cycle" intéressant de 10 années. Les chiffres sont de 114 fréquences annuelles, qu'on exprimera pour des raisons de commodité en logarithmes. A partir de ces données, on calcule les valeurs I_p (voir le tableau I dans l'appendice). A cause de la valeur très forte de $p = 12$, on ne peut déduire immédiatement s'il n'y a pas lieu d'introduire un élément spectral discret. Cependant, si nous comparons les sommes des carrés des résidus dans les deux hypothèses H_0 et H_1 du paragraphe 4, nous avons:

$$\sum_0 = 35,236 - 21,025 = 14,211 \qquad \text{(111 degrés de liberté)}$$

$$\sum_1 = 5,788 \times 111/109 = 5,894^{(1)} \qquad 1/2(111) \log e(14,211/5,894) = 48,9 > \log e^{99} = 4,60$$

Aussi, entre H_0 et H_1, c'est H_1 qu'il faut choisir. (Moran est arrivé à la même conclusion, sans cependant avoir établi un test définitif, par la seule inspection du corrélogramme). Les coefficients a_1 et a_2 de l'équation auto-régressive

$$X_t + a_1 X_{t-1} + a_2 X_{t-2} = Y_t$$

ou X_t est mesuré à partir de sa moyenne d'échantillon, sont $-1,4101$ et $0,77_{33}$

De la formule

$$f(\omega) = \left[\sigma_y^2 / \pi \sigma_x^2\right] \beta(\omega) \beta^*(\omega)$$

(1) Le chiffre 5,788 est obtenu à partir des 112 résidus actuels donnés par Moran après avoir ajusté un modèle autorégressif du 2° ordre; il a été ajusté pour compenser les 2 premiers résidus manquants.

où

$$1/\left[\beta(\omega)\beta^*(\omega)\right]=(1-a_2)^2+a_1^2+2a_1(1+a_2)\cos\omega+4a^2\cos^2\omega = b(\omega)$$

nous pouvons construire facilement T_q, dans l'hypothèse H_1.

$$T_q=\sum_{p=1}^{q} I_p\, b(\omega) \Big/ \sum_{p=1}^{\frac{1}{2}n} I_p\, \rho(\omega) . \quad (\omega=2\pi p/n)$$

I_0 n'intervient pas, car elle correspond à la moyenne (ici non nulle). La différence absolue la plus grande de T_q par rapport à $q/57$ est 0,184 (pour q= 30) qui dépasse juste le niveau significatif de P= 0,05[1], de $1{,}36/\sqrt{57}$ = 0,180.

Dans cet exemple l'effet dominant de la fréquence au voisinage de p = 12 est, ainsi qu'on l'a noté au § 3 perdu dans le test T_q, et l'emploi du test U_q non pondéré pourrait paraître plus approprié. Pour ce test alternatif, comme on cherche à tester une $f(\omega)$ particulière plutôt qu'à trouver une bande générale de confiance, il est préférable de calculer

$$v=\frac{(2\pi)^2}{2n}\sum_{p=0}^{\frac{1}{2}n} f^2(\omega)=\frac{2}{n}\left(\frac{\sigma_Y^2}{\sigma_X^2}\right)^2\sum_{p=0}^{\frac{1}{2}n} 1/b^2(\omega)$$

à partir des valeurs <u>théoriques</u> de $b(\omega)$, particulièrement en vue d'obtenir la valeur d'échantillon maximum de I_p pour p = 12.

Avec σ_Y^2/σ_X^2 = 0,147783 (et la sommation comprend la moitié de la contribution de p= 0 et de p=$\frac{1}{2}$ n, comme pour $f(\omega)$)nous obtenons v = 4,430. Alternativement, on peut employer la formule équivalente

$$v \sim \pi\int_0^{\pi} f^2(\omega)\,d\omega=\sum_{-\infty}^{\infty}\rho_s^2$$

qui donne la vérification suffisante, 4,417.

Une différence significative P=0,05 de U_q (disposée dans le tableau I)à $F(\omega)$devra être de l'ordre de $1{,}36 \times \sqrt{4{,}43}/\sqrt{57}$ = 0,38 et malgré le grand saut de U_q pour p=12, cette valeur n'est pas atteinte. Bien plus, la contribution dominante de I_{12} souligne combien toute théorie asymptotique est inadéquate pour U_q (ou V_q) dans cet exemple[2].

---

(1) Voir figure.

(2) En utilisant alternativement la Statistique modifiée suggérée à la note de la page , avec la fonction $f(\omega)$ utilisée dans T_q donnant la fonction spectrale provisoire, nous calculons:

$$\frac{1}{4}(114)\sum\left[I_p/f(\omega_p)\right]^2\Big/\left[\sum I_p/f(\omega_p)\right]^2 = 1{,}0304$$

de telle sorte que maintenant $\sqrt{v}$ = 1,015. Comme on peut s'y attendre, cette valeur est voisine de l'unité, et montre qu'une bande de confiance approchée pour la vraie fonction $f(\omega)$ peut être obtenue en remplaçant les frontières parallèles a' de la distribution théorique cumulative 2 q/n du graphique T_q, par de nouvelles frontières déduites par des translations de la distribution cumulative observée T_q sur cette échelle provisoire.

Alors qu'on a eu une différence significative avec le test T_q, une comparaison de ces conclusions avec celles tirées des autres tests sur les corrélogrammes montre les limites des tests ci-dessus quand on les emploie conjointement avec une fonction f (ω) "pas très uniforme". (cf. mes remarques à la fin du § 3). Moran trouve que les coefficients de régression partielle du quatrième ordre et du cinquième ordre dépassent le niveau significatif P=0,05. Un calcul plus compréhensif des 27 premières autocorrélations d'échantillons (qui ont été données par Moran dans cette note) sur la base du test goodness of fit de Quenouille (1947), donne pour χ^2 total la forte valeur de 54,4 avec 25 degrés de liberté (1). Cette approximation par le corrélogramme semble la plus naturelle quand une hypothèse de structure autorégressive est sous examen. Néanmoins, quoique les tests des spectres ne soient pas aussi appropriés dans cet exemple, l'emploi du χ^2 du périodogramme apparaît comme suffisamment effectif.

Les nombres testés sont dans le tableau II; le groupement de quelques degrés de liberté a été décidé a priori par rapport aux contributions à la fonction de distribution théorique F (ω) et non par rapport aux valeurs I_p.

Tableau II

p	$I_p b(\omega)$	D.F.	p	$I_p b(\omega)$	D.F.
1,2	0.14013	4	12	0.41208	2
3,4	0.40526	4	13	0.01589	2
5,6	0.05908	4	14	0.00179	2
7	0.02310	2	15,16	0.02989	4
8	0.07789	2	17,18,19	0.04600	6
9	0.01371	2	20,28	1.15438	18
10	0.05135	2	29,57	4.87392	57
11	0.01707	2	TOTAL:	7.32154	113

Le test d'homogénéité de ces sommes (pondérées) I_p donne

$$C\chi^2 = 37{,}70 \quad \text{où} \quad C = 1{,}1246$$

c'est-à-dire un χ^2 corrigé de 33,52 (avec 14 degrés de liberté) significatif au niveau P = 0,01. Ce résultat nous montre à partir du périodogramme avec une évidence plus convaincante combien l'hypothèse autorégressive est inadéquate.

On doit remarquer que, tandis que la valeur pondérée de I_p pour p = 12 n'apparaît plus particulièrement anormale, il indique sur l'échelle originale une hauteur suffisante et un maxi-

(1) Il y a une certaine évidence théorique (Bartlett et Rajalakshmann, 1953) que ce test exagère le niveau significatif véritable pour n borné. Cependant un calcul du prochain terme dans l'expansion asymptotique pour la variance d'échantillon, dans notre hypothèse que celui-là est comparable au terme correspondant du χ^2 basé sur les corrélations directes des résidus (Bartlett et Diananda, 1950) donne un "χ^2 réduit" de 38,5 avec 18 degrés de liberté, encore significatif pour le niveau P = 0,01.

mum assez proche pour rendre indésirable une régularisation excessive du périodogramme (avec seulement 114 observations); aucune régularisation, d'après les méthodes du § 2 n'a été effectuée. Il serait évidemment possible de chercher à trouver des fonctions de densité spectrale f (ω) plus compatibles avec les I_p observées que celle que l'on a testée: mais mon premier but a été d'avancer et de donner des exemples de divers tests, plutôt que de faire un examen approfondi de ces données particulières.

Remerciements - A M. A.M. Walker pour ses critiques au sujet de ces Conférences. A Mrs. A. Linnert pour l'analyse du périodogramme du tableau I, et à Mademoiselle Barbara Appleby qui m'a aidée dans l'analyse du corrélogramme. A MM. Adikhari et Joshi pour leur traduction.

BIBLIOGRAPHIE

Bartlett, M.S. (1937) "Properties of sufficiency and statistical test", Proc. Roy. Soc. Lond. A. 160, 268.

Bartlett (1946) "On the theoretical specification and sampling properties of autocorrelated time-series", J. Roy Statistical Soc. (Suppl.) 7, 211.

(1949) "Some evolutionary stochastic processes", J. Roy. Statist. Soc. (B), 11, 211.

(1950) "Periodogram analysis and continuous spectra" Biometrika, 37, 1.

(1954) An Introduction to Schochastic Processes with special reference to methods and applications (Cambridge, in Press).

Bartlett, M.S. and Diananda, P.H. (1950) "Extensions of Quenouille"s test for autoregressive schemes J. Roy. Statist. Soc. (B) 12, 108.

Bartlett, M.S., and Rajalaskshman, D.S. (1953) "Goodness of fit tests for simultaneous autoregressive series", J. Roy. Statist. Soc. (B), 15.

Davenport, Jr., W.B., Johnson, R. A., and Middleton, D. (1952) "Statistical errors in measurements on random time functions", J. App. Phys. 23, 377.

Fisher, R.A. (1929) "Test of significance in harmonic analysis", Proc. Roy. Soc. Lond. A 125, 54.

Grenander, U (1950) "Stochastic processes and statistical inference", J. App. Phys. 23, 377. Ark.Mat. 1, 195.

(1951) "On empirical spectral analysis of stochastic processes", Ark. Mat. 1, 503.

Grenander, U., and Rosenblatt, M. (1952) "On spectral analysis of stationary time-series", Proc. Nat. Acad. Sci. 38, 519.

Grenander, U., and Rosenblatt, M. (1953) "Statistical spectral analysis of time series arising from stationary stochastic processes", Ann. Mat.Statist. 24, 537.

Kendall, M.G. (1946) Researches in Oscillatory Time-Series, Cambridge.

Kolmogorov A. (1953)"Sulla determinazione empirica di una legge di distribuzione", Inst. Ital. Attuari Giom. 4, 1

Levy, P. (1948) Processus Stochastiques et Mouvement Browninen (Paris).

Mann, H.B., and Wald, A. (1943)"On the statistical treatment of linear stochastic difference equations", Econometrica, 11, 173.

Moran, P.A.P. (1953)"The Statistical analysis of the Canadian lynx cycle. I. Structure and prediction", Austral J. Zool. 1, 163.

Quenouille, M.H. (1947)"A large-sample test for the goodness of fit of autoregressive schemes" J. Roy Statist.Soc 110, 123.

Rao, U.R. (1951)"The applicability of large sample tests for moving average and autoregressive schemes to series of short lenght-an experimental study.Part.3: The discriminant function approach in the classification of time series", Sankhyä, 11, 257.

Sargan, J.D. (1953)"An approximate treatment of the properties of the correlogram and periodogram" J.Roy.Statist Soc. (B), 15, 140.

Slutsky, E. (1927)"The summation of random causes as the source of cyclic processes", Problems of Economic Conditions, ed. by the conjuncture Institute, Moscow,3 Ne I. (later reprinted in Econometrica, 5, 105).

Tukey, J.W. (1949)"Measuring noise colour", (unpublished memorandum).

Walker, A.M.(1952)"Some properties of the asymptotic power functions of goodness-of-fit tests for linear autoregressive schemes", J.Roy Statist.Soc.(B), 14,117.

Whittle P. (1951) Hypothesis testing in time series analysis Uppsala).

(1952a)"Tests of fit in time series" Biometrika 39 309.

(1952b)"The simultaneous estimation of a time series harmonic components and covariance structure", Trabajos Estadistica, 3, 43.

(1953)"The analysis of multiple stationary time series, J.Roy.Statist.Soc. (B) 15, 125.

Wold, H. (1938) Analysis of stationary Time-Series Uppsala.

(1949)"A large-sample test of moving averages", J. Roy Statist. Soc. (B), 11, 297.

Yule, G.U. (1927)"On the method of investigating periodicities in disturbed series, with special reference to Wolfer's sun-spotnumbers". Philos.Trans.A,226,267

Tableau I

p,q	I_p	U_q	$I_p\,b(\omega)$	T_q	F (ω)	ΔF (ω)
0		0	0	0	0,00981	0,00981
1	0,3818	0,0108	0,05116	0,0070	0,02943	0,01962
2	0,7061	0,0309	0,08897	0,0192	0,04999	0,02056
3	2,2847	0,0957	0,27188	0,0563	0,07175	0,02176
4	1,2825	0,1321	0,13338	0,0745	0,09666	0,02491
5	0,4003	0,1435	0,03659	0,0795	0,12499	0,02833
6	0,2928	0,1518	0,02249	0,0826	0,15872	0,03373
7	0,3756	0,1625	0,02310	0,0858	0,20083	0,04211
8	1,6715	0,2099	0,07789	0,0964	0,25641	0,05558
9	0,4069	0,2215	0,01371	0,0983	0,33327	0,07686
10	2,1132	0,2814	0,05135	0,1053	0,43985	0,10658
11	0,9126	0,3074	0,01707	0,1076	0,57836	0,13851
12	21,0246	0,9041	0,41208	0,1639	0,71050	0,13214
13	0,5778	0,9205	0,01589	0,1661	0,80468	0,09418
14	0,0406	0,9217	0,00179	0,1664	0,86328	0,05860
15	0,1812	0,9268	0,01292	0,1681	0,89960	0,03632
16	0,1557	0,9312	0,01697	0,1704	0,92336	0,02376
17	0,0924	0,9339	0,01515	0,1725	0,93916	0,01580
18	0,0080	0,9341	0,00183	0,1728	0,95047	0,01131
19	0,0927	0,9367	0,02902	0,1767	0,95874	0,00827
20	0,0221	0,9374	0,00906	0,1780	0,96506	0,00632
21	0,2510	0,9445	0,13278	0,1961	0,96996	0,00490
22	0,3611	0,9547	0,24049	0,2290	0,97385	0,00389
23	0,1777	0,9598	0,14589	0,2489	0,97700	0,00315
24	0,3175	0,9688	0,31814	0,2924	0,97959	0,00259
25	0,0339	0,9698	0,04051	0,2979	0,98175	0,00216
26	0,0325	0,9707	0,04596	0,3042	0,98358	0,00183
27	0,1054	0,9737	0,17359	0,3279	0,98516	0,00158
28	0,0252	0,9744	0,04796	0,3345	0,98652	0,00136
29	0,00013	0,97438	0,00022	0,3345	0,987703	0,001187
30	0,02294	0,97503	0,05672	0,3422	0,988749	0,001046
31	0,05902	0,97671	0,16426	0,3647	0,989679	0,000930
32	0,04530	0,97800	0,14111	0,3840	0,990510	0,000831
33	0,16174	0,98259	0,55819	0,4608	0,991261	0,000751
34	0,13346	0,98637	0,50877	0,5298	0,991940	0,000679
35	0,02145	0,98698	0,08978	0,5420	0,992559	0,000619
36	0,01871	0,98751	0,08510	0,5536	0,993129	0,000570
37	0,06197	0,98927	0,30578	0,5954	0,993654	0,000525
38	0,00266	0,98935	0,01435	0,5974	0,994141	0,000487
39	0,02603	0,99009	0,14823	0,6176	0,994595	0,000454
40	0,02213	0,99072	0,13450	0,6360	0,995019	0,000424
41	0,04949	0,99212	0,32031	0,6798	0,995421	0,000402
42	0,01225	0,99247	0,08421	0,6913	0,995799	0,000378
43	0,04148	0,99365	0,29938	0,7322	0,996159	0,000360
44	0,04151	0,99482	0,31418	0,7751	0,996501	0,000342
45	0,02565	0,99555	0,20336	0,8029	0,996827	0,000326
46	0,01052	0,99585	0,08651	0,8147	0,997141	0,000314
47	0,04077	0,99701	0,34866	0,8624	0,997444	0,000303
48	0,00668	0,99720	0,05917	0,8705	0,997736	0,000292
49	0,00369	0,99730	0,03664	0,8751	0,998021	0,000285
50	0,01218	0,99765	0,11581	0,8906	0,998298	0,000277
51	0,01648	0,99812	0,15737	0,9121	0,998570	0,000272
52	0,03090	0,99899	0,30026	0,9531	0,998837	0,000267
53	0,00150	0,99904	0,01480	0,9552	0,999099	0,000262
54	0,01879	0,99957	0,18764	0,9808	0,999358	0,000259
55	0,01051	0,99987	0,10571	0,9952	0,999614	0,000256
56	0,00325	0,99996	0,03338	0,9998	0,999870	0,000256
57	0,00013	1,0000	0,00152	1,0000	1,000000	0,000130
TOTAL:	35,230		7,32154			1,000000

SUR UNE ESTIMATION EXHAUSTIVE POUR LA MOYENNE D'UNE VARIABLE ALÉATOIRE OBÉISSANT A LA LOI DE LAPLACE DANS UN ESPACE DE BANACH

par

Shafik DOSS

Introduction. Soit x une variable aléatoire prenant ses valeurs dans un espace de Banach B. On dira II (1) que x obéit à une loi laplacienne si, quelle que soit la fonctionnelle linéaire f, définie dans B, la variable aléatoire numérique $f(x) = x_f$ obéit à la loi de Laplace.

Observons que si l'on se donne la distribution de x_f pour toutes les fonctionnelles f, on a aussi, quel que soit n, la distribution de la variable $(x_{f_1}, x_{f_2}, \dots, x_{f_n}) \in R^n$ [V]. En effet, considérons la fonctionnelle $t_1 f_1 + t_2 f_2 + \dots + t_n f_n$ où les t_i sont des nombres réels. On a

$$x_{t_1 f_1 + t_2 f_2 + \dots + t_n f_n} = t_1 x_{f_1} + t_2 x_{f_2} + \dots + t_n x_{f_n}$$

La valeur moyenne

$$M(e^{i x_{t_1 f_1 + \dots t_n f_n}}) = M(e^{i \sum_{s=1}^{n} t_s x_{f_s}})$$

est connue, par hypothèse, quels que soient les t_s ; mais ce n'est autre que la fonction caractéristique de la variable cartésienne $(x_{f_1}, x_{f_2}, \dots, x_{f_n})$. D'après un théorème classique (par ex. [I] p. 101), la distribution de la variable $(x_{f_1}, \dots, x_{f_n})$ est déterminée.

Remarquons encore que la fonction caractéristique d'une variable numérique laplacienne y étant $e^{itM(y) - \frac{1}{2}t^2 M(y - M(y))^2}$, on aura

$$M(e^{i \sum_{s=1}^{n} t_s x_{f_s}}) = e^{i \sum_s m_s t_s - \frac{1}{2} \sum_{r,s} \lambda_{rs} t_r t_s} \tag{1}$$

où l'on a posé $m_s = M(x_{f_s})$, $\lambda_{rs} = M(x_{f_r} - m_r)(x_{f_s} - m_s)$. Ici λ_{rs} existe pour tout couple (r,s) en vertu de l'inégalité de Schwarz puisque, x_{f_i} étant laplacienne, λ_{ii} existe pour tout i.

---

(1) Les nombres entre crochets renvoient à la Bibliographie à la fin du mémoire.

La formule (1) montre que la variable aléatoire $(x_{f_1}, x_{f_2}, \ldots, x_{f_n})$ obéit à la loi de Laplace définie par les moments du premier ordre m_s et les moments du second ordre λ_{rs}. On peut donc dire que la variable aléatoire $x \in B$ est laplacienne si, quels que soient n et les fonctionnelles $f_1, f_2, \ldots f_n$, la variable cartésienne $(x_{f_1}, x_{f_2}, \ldots, x_{f_n})$ est laplacienne.

Soient a_i, b_i des nombres réels satisfaisant à $-\infty \leqslant a_i < b_i \leqslant +\infty$ L'ensemble des valeurs x satisfaisant à

$$a_i < x_{f_i} < b_i \quad ; \quad i = 1,2, \ldots, n$$

sera appelé un intervalle I.

Nos définitions impliquent qu'une mesure est donnée sur un corps borélien l de sous-ensembles de B et que l'on a $I \in l$ pour tout intervalle I. Soit K le plus petit corps borélien contenant tous les I. On aura évidemment $K < l$. Nous considèrerons dans la suite exclusivement les ensembles de K. Il est bien connu [VI] que la mesure, définie pour tous les I, est déterminée de façon univoque pour tous les ensembles de K. Deux lois de probabilité qui coïncident sur les ensembles de K seront considérées comme identiques.

La donnée d'une loi laplacienne pour x est équivalente à la donnée des lois laplaciennes correspondantes pour les x_f et celles-ci sont équivalentes à la donnée des $m_f = M(x_f)$ et $\lambda_{fg} = M(x_f - m_f)(x_g - m_g)$.

Considérons la famille des lois laplaciennes pour lesquelles les λ_{fg} sont fixés et les m_f sont variables. La donnée des m_f est équivalente à la donnée de la moyenne M(x), quand elle existe, de la variable aléatoire x. Cette moyenne étant définie [IV] par l'égalité

$$f(M(x)) = m_f \quad \text{pour toutes les fonctionnelles } f .$$

Considérant n déterminations indépendantes $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ d'une variable aléatoire x obéissant à une loi de la famille considérée, et formant l'élément $\eta = (x^{(1)} + \ldots + x^{(n)})/n$, il est naturel, généralisant un résultat connu (voir par ex. [I] p.495, ex. 2), de présumer que η constitue une estimation exhaustive du paramètre M(x) qui détermine la loi considérée. C'est ce que nous allons établir. Mais rappelons d'abord la définition donnée par Halmos et Savage (III) des estimations exhaustives.

Définition de Halmos et Savage. Soient X et Y deux ensembles abstraits et soient S et $\mathcal{T}$ deux corps boréliens de sous-ensembles de X et Y respectivement. On dira que la fonction T définie sur X et prenant ses valeurs dans Y est mesurable si quel que soit $F \in \mathcal{T}$, on a $T^{-1}(F) \in S$. On supposera, dans ce qui suit, T mesurable et, si Y est un espace euclidien, on supposera que $\mathcal{T}$ sera composé des ensembles boréliens. Si g est une fonction définie sur Y, on désignera par gT la fonction définie sur X par l'égalité $gT(x) = g(T(x))$. Si μ est une mesure définie sur S, μT^{-1} sera la mesure définie sur $\mathcal{T}$ par l'égalité.. $\mu T^{-1}(F) = \mu(T^{-1}(F))$. On établit aisément l'égalité

$$(2) \qquad \int_F g(y)\, d\mu T^{-1}(y) = \int_{T^{-1}(F)} gT(x)\, d\mu(x).$$

Si μ et ν sont deux mesures définies sur S et si ν est absolument continue par rapport à μ, on aura, d'après le théorème de Radon-Nikodym

$$\nu(E) = \int_E f(x)\, d\mu(x)$$

pour tout $E \in S$. On écrira aussi cette relation

$$d\nu = f\, d\mu$$

Il est facile d'établir que si ν est absolument continue par rapport à μ, alors νT^{-1} est absolument continue par rapport à μT^{-1}. On aura donc

$$d\nu T^{-1} = g\, d\mu T^{-1}$$

g (y) sera appelée [VI], [III] l'espérance mathématique conditionnée de f (x) pour y fixé. On la désignera par $e_\mu(f/y)$. Si $f = X_E$ est la fonction caractéristique d'un ensemble $E \in S$, $e_\mu(f/y)$ sera appelée probabilité conditionnée de E pour y fixé. On la désignera par $p_\mu(E/y)$.

L'équation définissant $e_\mu(f/y)$ s'écrit

$$\int_{T^{-1}(F)} f(x)\, d\mu(x) = \int_F e_\mu(f/y)\, d\mu T^{-1}(y), \quad F \in \mathcal{C}.$$

et celle définissant $p_\mu(E/y)$ s'écrit

(3) $$\mu(E \cap T^{-1}(F)) = \int_F p_\mu(E/y)\, d\mu T^{-1}(y), \quad F \in \mathcal{C}.$$

Soit M une famille de mesures μ sur S. On dira que T est une estimation exhaustive pour M si, quel que soit $E \in S$, il existe une fonction mesurable $p = p(E/y)$ sur Y telle que pour tout $\mu \in M$

$$p_\mu(E/y) = p(E/y)$$

soit vérifiée sauf sur un ensemble de valeurs y de mesure μT^{-1} nulle. Cet ensemble de mesure nulle dépend évidemment de E.

Appliquons ces notions au cas qui nous occupe:

L'ensemble X sera B^n. Un point ξ de X sera $(x^{(1)}, x^{(2)}, \ldots, x^{(n)})$, où $x^{(i)} \in B$. L'ensemble Y sera B. S sera le plus petit corps boréLien contenant les ensembles $E_1 \times E_2 \times \ldots \times E_n$, où $E_i \in K$. $\mathcal{C}$ sera K. L'estimation T sera donnée par

$$\eta = T(\xi) = \frac{1}{n}(x^{(1)} + x^{(2)} + \ldots + x^{(n)}) \in Y.$$

La famille M de mesures μ sera celle qu'on obtient en prenant sur X les mesures produits des mesures définies par la donnée de (m_f) sur B ; c'est-à-dire que l'on posera $\mu(E_1 \times \ldots \times E_n) = g(E_1)\, g(E_2) \ldots g(E_n)$, où g désigne la mesure laplacienne définie sur les ensembles de K par la don-

née de (m_f) <u>variable</u> et des λ_{fg} <u>fixés</u>. L'estimation T est évidemment mesurable. Cela se voit en observant que $x_f^{(i)}$ est mesurable S sur X .

<u>L'estimation T est exhaustive</u>. Pour montrer que T est une estimation exhaustive pour la famille M de mesures μ, il faut montrer que, quel que soit $E \in S$, il existe une fonction mesurable $p = p\,(E/y)$ sur Y telle que, pour tout $\mu \in M$,

$$p_\mu\,(E/y) = p\,(E/y)$$

soit vérifiée, sauf sur un ensemble de valeurs η de mesure μT^{-1} nulle.

Soient E un ensemble de S_{s_0} où S_{s_0} est le corps borélien dans X composé des ensembles de S définis au moyen des s_0 fonctionnelles $f_{i_1}, f_{i_2}, \ldots, f_{i_{s_0}}$, et $F \in \mathcal{C}_s$ où $\mathcal{C}_s$ est le corps borélien dans Y composé des ensembles de $\mathcal{C}$ définis au moyen des s fonctionnelles $f_{i_1}, \ldots, f_{i_{s_0}}, f_{i_{s_0+1}}, \ldots, f_{i_s}$.

Il est clair que $E \cap T^{-1}(F)$ sera défini par des conditions sur les $x_{f_{i_\nu}}^{(r)}$; $r = 1, 2, \ldots, n$; $\nu = 1, 2, \ldots, s$. Dans l'espace euclidien R^{ns}, considérons le point ξ' dont les coordonnées sont $x_{f_{i_\nu}}^{(r)}$ et soit θ la transformation $\theta(\xi) = \xi'$. Il est clair que θ est mesurable et ξ' obéit à une loi laplacienne dans R^{ns}. On désignera la mesure pour cette loi laplacienne par μ'.

Soient $E' \subset R^{ns}$ défini par $E = \theta^{-1}(E')$ et G' défini par $T^{-1}(F) = \theta^{-1}(G')$. On aura évidemment

$$\mu\,(E \cap T^{-1}(F)) = \mu'(E' \cap G'). \tag{4}$$

Soient aussi ψ_s la transformation mesurable définit sur Y par

$$\rho_s = \psi_s(\eta) = (\eta_{f_{i_1}}, \eta_{f_{i_2}}, \ldots, \eta_{f_{i_s}}) \in R^s$$

et $C \subset R^s$ défini par $F = \psi_s^{-1}(C)$; et soit Λ la transformation mesurable définie sur R^{ns} par

$$\rho_s = \Lambda(\xi') = (\eta_{f_{i_1}}, \eta_{f_{i_2}}, \ldots, \eta_{f_{i_s}}) \in R^s$$

où

$$\eta_{f_{i_\nu}} = \frac{1}{n}\,(x_{f_{i_\nu}}^{(1)} + \ldots + x_{f_{i_\nu}}^{(n)})$$

On a évidemment

$$\psi_s T(\xi) = \Lambda\theta(\xi) \quad, \xi \in X \ ;$$

d'où l'on déduit

$$C = \psi_s \psi_s^{-1}(C) = \psi_s(F) = \psi_s T T^{-1}(F) =$$

$$= \Lambda\theta T^{-1}(F) = \Lambda\theta\theta^{-1}(G') = \Lambda(G')$$

On en tire

(5) $$G' = \Lambda^{-1}(C)$$

car il est clair que F étant défini au moyen des $\eta_{f_{i_\nu}}$ seulement G' l'est aussi.

Cela étant, on a en tenant compte de (4), (5), (3) et (2) et en observant que $\mu T^{-1} \psi_s^{-1}$ est la mesure sur R^s,

(6) $$\mu(E \cap T^{-1}(F)) = \mu'(E' \cap G') = \mu'(E' \cap \Lambda^{-1}(C)) =$$

$$= \int_C g_s(\rho_s)\, d\mu\, T^{-1} \psi_s^{-1}(\rho_s) = \int_F g_s \psi_s(\eta)\, d\mu\, T^{-1}(\eta)$$

où $g_s(\rho_s) = p_{\mu'}(E'/\rho_s)$ est indépendant de la mesure μ et par suite de μ car (I , p. 495) ρ_1 est une estimation exhaustive de la moyenne de la variable laplacienne $(x_{f_1}, x_{f_2}, \dots, x_{f_s})$. On tire de (6) en tenant compte de la formule générale (3),

$$\int_F p_\mu(E/\eta)\, d\mu T^{-1}(\eta) = \int_F g_s \psi(\eta)\, d\mu T^{-1}(\eta) \quad , F \in \mathcal{C}_s$$

Or, on va montrer que $g_s\ (\rho_s) = g_{s_0}(\rho_{s_0})$. On aura donc, quel que soit $\mathcal{C}_s, \mathcal{C}_{s_0}$

(7) $$\int_F p_\mu(E/\eta)\, d\mu T^{-1}(\eta) = \int_F g_{s_0} \psi_{s_0}(\eta)\, d\mu\, T^{-1}(\eta), \ F \in \mathcal{C}_s .$$

L'égalité dans (7) étant vraie pour tous les $F \in \mathcal{C}_1$ quel que soit $\mathcal{C}_1 \supset \mathcal{C}_{s_0}$, elle est aussi vraie pour tout $F \in \mathcal{C}$. Car, comme on le voit immédiatement, les ensembles $F \in \mathcal{C}$ pour lesquels l'égalité dans (7) est vraie, forment un corps borélien. Ce corps comprenant les ensembles de tous les $\mathcal{C}_s \supset \mathcal{C}_{s_0}$, il coïncide avec $\mathcal{C}$ .On a donc

(8) $$\int_F p_\mu(E/\eta)\, d\mu T^{-1}(\eta) = \int_F g_{s_0} \psi_{s_0}(\eta)\, d\mu T^{-1}(\eta) \quad , F \in \mathcal{C}$$

On tire de (8), par application du théorème de Radon-Nikodym, $p_\mu\ (E/\eta) = g_{s_0} \psi_{s_0}(\eta)$ presque partout.
$p_\mu\ (E/\eta)$ est donc indépendant de μ (p. partout).

La classe des ensembles $E \in S$ pour lesquels $p_\mu\ (E/\eta)$ est indépendant de μ (p. partout), forme évidemment un corps borélien Ce corps comprenant tous les ensembles de S_s quel que soit s, puisque s_0 était arbitraire; il coïncide avec S . Car il est clair que S est le plus petit corps borélien comprenant tous les ensembles de S_s quel que soit s .

Ceci achève la démonstration.

Une propriété de la loi de Laplace dans un espace euclidien. Revenons au point laissé en suspens. Il s'agit de montrer que

$$g_s \psi_s(\eta) = g_{s_0} \psi_{s_0}(\eta) \quad \text{pour } \mathcal{C}_s, \mathcal{C}_{s_0} .$$

Considérons l'ensemble E'. Il est défini par des conditions sur les $x^{(r)}_{f_{i_\tau}}$, $r = 1, 2, \ldots, n$; $\tau = 1, 2, \ldots, s_0$; et l'on a

$$g_s \psi_s(\eta) = g_s(\rho_s) = p\,(E'/\eta_{f_{i_1}}; \eta_{f_{i_2}}, \ldots, \eta_{f_{i_s}})$$

$$g_{s_0} \psi_{s_0}(\eta) = g_{s_0}(\rho_{s_0}) = p\,(E'/\eta_{f_{i_1}}, \eta_{f_{i_2}}, \ldots, \eta_{f_{i_{s_0}}})$$

où

$$\eta_{f_{i_\nu}} = \frac{1}{n}(x^{(1)}_{f_{i_\nu}} + \ldots + x^{(n)}_{f_{i_\nu}}).$$

Changeons les notations. Le problème sera comme suit :

Soit $x=(x_1, x_2, \ldots, x_r)$ une variable dans R^r et $y=(y_1, y_2, \ldots, y_t)$ une variable dans R^t. On suppose que (x, y) est une variable à loi laplacienne non dégénérée dans R^{r+t} ; alors, x et y seront des variables à lois laplaciennes non dégénérées dans R^r et R^t respectivement.

On considère n déterminations indépendantes $(x^{(i)}, y^{(i)})$, $i = 1, 2, \ldots, n$, de la variable (x,y) dans R^{r+t} et un ensemble borélien $E \subset R^{rn}$. Posons $\xi=(x^{(i)}_u)$, $\eta=(y^{(i)}_v)$, $1 \leq i \leq n$; $1 \leq u \leq r$; $1 \leq v \leq t$. On aura $\xi \in R^{rn}$, $\eta \in R^{tn}$, $(\xi, \eta) \in R^{(r+t)n}$.

On forme $\alpha_u = \frac{1}{n}(x^{(1)}_u + \ldots + x^{(n)}_u)$, $\beta_v = \frac{1}{n}(y^{(1)}_v + \ldots + y^{(n)}_v)$ et on considère les probabilités conditionnelles

$$p\,(\xi \in E/\alpha_1, \alpha_2, \ldots, \alpha_r ; \beta_1, \beta_2, \ldots, \beta_t)$$

et

$$p\,(\xi \in E/\alpha_1, \alpha_2, \ldots, \alpha_r).$$

Il s'agit de montrer que <u>ces probabilités conditionnelles sont égales p. partout,</u> c'est-à-dire égales sauf un cas de probabilité nulle.

Introduisons des variables z_i par les équations

$$z_i = y_i + k^{(1)}_i x_1 + k^{(2)}_i x_2 + \ldots + k^{(r)}_i x_r \quad , i = 1,2,\ldots,t$$

et déterminons les constantes $k^{(j)}_i$ par les conditions que les variables z_i soient indépendantes des x_j. Il suffira d'écrire

(9) $M\,(z_i - Mz_i)\,(x_j - Mx_j) = 0$, $i = 1,2,\ldots,t$; $j = 1,2,\ldots, z_r$

Les conditions (9) détermineront d'une façon univoque les $k^{(j)}_i$ car la loi laplacienne de (x,y) est non dégénérée. La variable $z = (z_1, z_2, \ldots, z_t)$ obéira à une loi laplacienne indépendante de la loi de $x = (x_1, x_2, \ldots, x_r)$. Soient $(x^{(i)}, z^{(i)})$ les n déterminations indépendantes de (x, z) correspondant aux n déterminations indépendantes $(x^{(i)}, y^{(i)})$ de (x, y). On pose

$$\gamma_j = \frac{1}{n}(z^{(1)}_j + \ldots + z^{(n)}_j) = \beta_j + k^{(1)}_j \alpha_1 + k^{(2)}_j \alpha_2 + \ldots + k^{(r)}_j \alpha_r \ , \ j=1,2,\ldots,t$$

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_r), \quad \beta = (\beta_1, \beta_2, \ldots, \beta_t), \quad \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_t)$$

Désignons par λ la propabilité pour la variable (α,β), par π la probabilité pour la variable (α,γ) dans R^{r+t} et soit T la correspondance biunivoque $T(\alpha,\beta) = (\alpha,\gamma)$. On a $\pi = \lambda T^{-1}$ et, si G est un ensemble borélien quelconque dans R^{r+t},

$$p(\xi \in E \cap (\alpha,\beta) \in G) = \int_G p_1(\alpha,\beta)\, d\lambda(\alpha,\beta) = \int_{T(G)} p_1 T^{-1}(\alpha,\gamma)\, d\pi(\alpha,\gamma)$$

où l'on a posé $p_1(\alpha,\beta) = p(\xi \in E/(\alpha,\beta))$. De plus.

$$(10)\quad p(\xi \in E \cap (\alpha,\beta) \in G) = p(\xi \in E \cap (\alpha,\gamma) \in T(G)) = \int_{T(G)} p(\xi \in E/(\alpha,\gamma))\, d\pi(\alpha,\gamma).$$

On en tire

$$p(\xi \in E/(\alpha,\gamma)) = p_1 T^{-1}(\alpha,\gamma) = p_1(\alpha,\beta) = p(\xi \in E/(\alpha,\beta)) \qquad \text{p. partout}$$

Or, comme il est facile de le voir, $p(\xi \in E/(\alpha,\gamma))$ est p. partout indépendant de γ et, par suite, $p(\xi \in E/(\alpha,\beta))$ est p. partout indépendant de β.

Pour voir que $p(\xi \in E/(\alpha,\beta))$ ne dépend p. partout que de α, observons d'abord que α et γ sont deux variables laplaciennes indépendantes. Soient Φ la mesure pour α et ψ la mesure pour γ, π étant, comme ci-haut, la mesure pour (α,γ).

Soient F un ensemble borélien de R^r et G un ensemble borélien de R^t. En observant que l'évènement $\gamma \in G$ est indépendant des évènements $\xi \in E$, $\alpha \in F$, on obtient

$$(11)\quad \int_{F\times G} p(\xi \in E/(\alpha,\gamma))\, d\pi(\alpha,\gamma) = p(\xi \in E \cap (\alpha,\gamma) \in F\times G) = p(\xi \in E \cap \alpha \in F \cap \gamma \in G) =$$

$$= p(\xi \in E \cap \alpha \in F)\, p(\gamma \in G) = \int_G d\psi(\gamma) \int_F p(\xi \in E/\alpha)\, d\Phi(\alpha) = \int_{F\times G} p(\xi \in E/\alpha)\, d\pi(\alpha,\gamma),$$

d'après le théorème de Fubini. Or, les ensembles $Q \subset R^{r+t}$ pour lesquels

$$\int_Q p(\xi \in E/(\alpha,\gamma))\, d\pi(\alpha,\gamma) = \int_Q p(\xi \in E/\alpha)\, d\pi(\alpha,\gamma)$$

forment un corps. Ce corps comprenant, d'après (11), tous les ensembles $F \times G$ où $F \subset R^r$ et $G \subset R^t$ sont boréliens; il comprend tous les ensembles boréliens de R^{r+t}. Le théorème de Radon-Nikodym donne alors en tenant compte de (10)

$$p(\xi \in E/(\alpha,\beta)) = p(\xi \in E/(\alpha,\gamma)) = p(\xi \in E/\alpha) \quad \text{p. partout.}$$

Ceci achève la démonstration.

BIBLIOGRAPHIE

[I] Cramér, H.; Mathematical Methods of Statistics, Princeton (1946).

[II] Fréchet, M.; Généralisations de la loi de probabilité de Laplace, Ann. Inst. H.Poincaré t.XII, fasc.I (1951)p.1-29

[III] Halmos, P.R. et Savage, L.J.; Application of the Radon-Nikodym Theorem to the Theory of Sufficient Statistics, Ann. Math. Stat. vol. XX, no.2 (1949, pp.225-241.

[IV] Mourier, E.; Sur l'espérance mathématique d'un élément aléatoire dans un espace de Banach, C.R. Acad. Sci. Paris t. 229 (1949) p. 1300-1301.

[V] Mourier, E.; Propriétés des caractéristiques d'un élément aléatoire dans un espace de Banach, C.R. Acad. Sci. Paris t. 231 (1950) p. 28-29.

[VI] Kolmogoroff, A.; Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin, Springer (1933).

SUR LE THÉORÈME LIMITE CENTRAL POUR DES VARIABLES ALÉATOIRES DANS UN ESPACE DE BANACH

par

Shafik DOSS

Dans un espace de Banach B, on dira que la variable aléatoire x est laplacienne si, pour toute fonctionnelle linéaire f, la variable aléatoire numérique f (x) est laplacienne [VII](1). Il revient au même de dire que pour toute combinaison finie de fonctionnelles f_1, f_2, ..., f_n, la variable aléatoire $(f_1(x), f_2(x), \ldots, f_n(x)) \in R^n$ est laplacienne [VI]

L'objet du présent travail est de montrer que, sous certaines conditions, si (x_n) est une suite de variables aléatoires indépendantes dans l'espace de Banach B, la variable $y_n = (x_1 + \ldots + x_n)/\sqrt{n}$ converge, au sens de Bernoulli, vers une variable laplacienne. C'est l'extension, au cas qui nous occupe, du théorème limite central du calcul des probabilités.

"Un cas particulier de ce théorème a été étudié par E. Mourier (IX). C'est le cas où B est un espace de Hilbert séparable et où les x_i ont la même distribution avec $m(\|x_i\|) < \infty$ $(i = 1,2,\ldots,n)$. Un cas un peu plus général a aussi été considéré par Fortet et Mourier (X)".

Introduction. Soit B un espace de Banach, l un corps borélien de sous-ensembles de B et μ une mesure (probabilité) sur les ensembles de l . Soient f_1, f_2, ..., f_n des fonctionnelles linéaires quelconques définies sur B et a_i, b_i, $-\infty \leqslant a_i < b_i \leqslant +\infty$, des nombres quelconques. L'ensemble des valeurs de $x \in B$ pour lesquelles

$$(1) \qquad a_i < f_i(x) < b_i \quad , i = 1, 2, \ldots, n ,$$

avec le signe d'égalité pour $a_i = -\infty$ et pour $b_i = +\infty$, sera appelé un intervalle $\mathcal{L}$ de B. On supposera que tous les intervalles appartiennent à l et on considèrera le plus petit corps borélien K contenant tous les $\mathcal{L}$. La mesure μ sera considérée sur K . Elle est complètement déterminée par sa donnée sur les intervalles $\mathcal{L}$ [VIII] .

(1) Les nombres entre crochets renvoient à la Bibliographie à la fin du mémoire.

Se donner une variable aléatoire x sur B ne sera pas autre chose que se donner la mesure μ pour les ensembles de K.

Soient $x_1, x_2, \ldots, x_n$ des variables aléatoires indépendantes définies sur B. Cela signifie que dans l'espace B^n des points $(x_1, x_2, \ldots, x_n)$, les ensembles de la forme $\mathcal{L}_1 \times \mathcal{L}_2 \times \ldots \times \mathcal{L}_n$ ont pour mesure $\mu^{(n)}(\mathcal{L}_1 \times \ldots \times \mathcal{L}_n) = \mu_1(\mathcal{L}_1)\,\mu_2(\mathcal{L}_2)\ldots\mu_n(\mathcal{L}_n)$. La mesure dans B^n sera alors définie de façon univoque sur le plus petit corps borélien $K^{(n)}$ contenant tous les ensembles de la forme $\mathcal{L}_1 \times \mathcal{L}_2 \times \ldots \times \mathcal{L}_n$.

Soit T la transformation $y_n = T(x_1, \ldots, x_n) = \frac{1}{\sqrt{n}}(x_1 + \ldots + x_n) \in B$. Il est clair que T est mesurable en ce sens que si $\alpha \in K$, on a $T^{-1}(\alpha) \in K^{(n)}$. On posera pour la variable aléatoire y_n,

$$\mu_n(\alpha) = \mu^{(n)}(T^{-1}(\alpha)).$$

Ainsi, la donnée des variables aléatoires indépendantes $x_1, \ldots, x_n$ détermine la mesure μ_n sur les ensembles de K, c'est-à-dire la variable aléatoire $y_n = \frac{x_1 + \ldots + x_n}{\sqrt{n}}$.

Les espaces Ω et Ω_1 d'après le modèle de Doob. B étant donné l'espace conjugué B* dont les éléments sont les fonctionnelles linéaires f définies sur B est déterminé. Considérons l'ensemble Ω des points $(x_f)_{f \in B^*}$, $-\infty \leqslant x_f \leqslant +\infty$ et le sous-ensemble Ω_1 composé des points (x_f) tels qu'il existe $x \in B$ pour lequel

$$f(x) = x_f \quad \text{pour tout} \quad f .$$

Il est clair que la correspondance, ainsi établie, entre les points de B et ceux de Ω_1 est biunivoque.

Dans la suite, Ω_1 remplacera B et il sera toujours considéré comme plongé dans Ω.

Un intervalle L de Ω sera l'ensemble des points de Ω qui satisfont à des conditions de la forme

$$(2) \qquad a_i < x_{f_i} < b_i \quad ; \quad i = 1, 2, \ldots, n \quad ; \quad -\infty \leqslant a_i < b_i \leqslant +\infty$$

avec le signe d'égalité pour $a_i = -\infty$ et pour $b_i = +\infty$. Tout intervalle sera considéré comme un voisinage dans Ω de l'ensemble des points qui lui appartiennent. Ω devient ainsi un espace topologique compact et les intervalles forment, par leur définition même, un système fondamental de voisinages ouverts [IV].

On sait que tout espace compact est susceptible d'une structure uniforme et d'une seule [I]. Nous considérerons Ω comme espace uniforme et les entourages V qui définissent la structure uniforme seront pris ouverts. Ω_1 sera muni de la structure uniforme induite par celle de Ω. Un intervalle $L^{(1)}$ de Ω_1 sera tout ensemble de la forme $L \cap \Omega_1$.

B étant muni de la mesure μ sur K ; à tout intervalle $\mathcal{L}$ de B défini par (1), il correspond un intervalle L de Ω défini par (2) et un intervalle $L^{(1)}$ de Ω_1. On définira la mesure m de L par

$$(3) \qquad m(L) = \mu(\mathcal{L})$$

pour toute fonctionnelle linéaire f. Supposons aussi que les quantités

$$\lambda_{rs}^{(n)} = M f_r (x_n) \; f_s \; (x_n)$$

existent pour tout n et pour tout couple f_r, f_s de fonctionnelles linéaires. On supposera de plus que pour tout groupe fini de fonctionnelles f_1, f_2, ..., f_k, les conditions suivantes sont satisfaites

(i) $$\lim_{n\to\infty} \frac{1}{n} \sum_{\tau=1}^{n} \lambda_{rs}^{(\tau)} = \lambda_{rs} \qquad (r,s = 1,2,\ldots,k)$$

où les λ_{rs} ne sont pas tous nuls

(ii) $$\lim_{n\to\infty} \frac{1}{n} \sum_{\tau=1}^{n} \int_{|X|>\varepsilon\sqrt{n}} |X|^2 \, d\, P_\tau = 0 \quad \text{pour tout } \varepsilon > 0$$

où P_τ est la distribution de la variable $(f_1(x_\tau),\ldots, f_k(x_\tau)) \epsilon R^k$

et $$|X| = \sqrt{[f_1(x_\tau)]^2 + \ldots + [f_k(x_\tau)]^2}$$

Alors, d'après un théorème connu ([II], p. 113), la loi de distribution de la variable $(f_1(y_n), \ldots, f_k(y_n)) \in R^k$ converge vers la loi laplacienne dont les premiers moments sont nuls et les seconds égaux à λ_{rs}.

Passant à Ω et Ω_1 ; si l'on désigne par g la mesure laplacienne qu'on vient de décrire, cette mesure sera définie sur tous les intervalles L de Ω et, d'après ce qu'on vient de montrer, on a

$$\lim_{n\to\infty} m_n \; (L) = g \; (L) \quad \text{pour tout intervalle } L.$$

Pour que la mesure g puisse être définie sur les ensembles du corps $K^{(1)}$ dans Ω_1, il faudra, d'après la Remarque, que

$$\Omega_1 \subset \bigcup_i L_i \quad \text{implique} \quad \lim_{n\to\infty} g\,(\bigcup_{i=1}^{n} L_i) = 1 .$$

Nous supposerons, dans la suite, cette condition satisfaite. C'est une condition sur les λ_{rs}. On peut alors poser

$$g^{(1)} \; (A^{(1)}) = g \; (A) \quad \text{pour} \quad A_1 = A \cap \Omega_1 \epsilon K^{(1)}$$

et l'on aura

$$\lim_{n\to\infty} m_n^{(1)}(L^{(1)}) = g^{(1)} \; (L^{(1)}) \text{ pour tout intervalle } L^{(1)} \text{ de } \Omega_1 .$$

Pour pouvoir affirmer que la loi $m_n^{(1)}$ converge, au sens de Bernoulli, vers la loi $g^{(1)}$, il faut montrer que pour tout ensemble de continuité $A^{(1)}$ pour $g^{(1)}$ [V], on a

$$\lim_{n\to\infty} m_n^{(1)} \; (A^{(1)}) = g^{(1)} \; (A^{(1)}) .$$

La mesure $m(L)$, ainsi définie, satisfait aux conditions de comptabilité énoncées par Kolmogoroff [VIII] et elle peut donc être étendue aux ensembles du plus petit corps borélien K contenant tous les intervalles L . Il est facile de voir [III] que Ω_1 à une mesure extérieure égale à 1. On peut alors pour tout ensemble $A^{(1)}$ de la forme $A \cap \Omega_1$ où $A \in K$ poser

$$m^{(1)}(A^{(1)}) = m(A) \tag{4}$$

et si $\mathfrak{A}$ est l'ensemble correspondant à $A^{(1)}$ dans la correspondance biunivoque établie entre B et Ω_1 , on aura

$$m^{(1)}(A^{(1)}) = \mu(\mathfrak{A})$$

et $m^{(1)}$ sera définie sur les ensembles du corps borélien $K^{(1)}$ composé des ensembles $A^{(1)}$. Ce corps $K^{(1)}$ correspond au corps K dans la correspondance entre B et Ω_1.

Remarque. Si μ n'est pas initialement donnée sur l , mais seulement sur les intervalles $\mathfrak{L}$ de B ; alors on peut toujours poser l'égalité (3) et, si les conditions de Kolmogoroff sont satisfaites pour m , on peut étendre la mesure m aux ensembles du corps K . Mais pour que l'on puisse écrire (4), c'est-à-dire étendre $m^{(1)}$ au corps $K^{(1)}$ ou, ce qui revient au même, étendre μ au corps K , il faut et il suffit que la mesure extérieure de Ω_1 soit égale à 1, c'est-à-dire que l'on ait $m(\bigcup_i L_i) = 1$ pour toute suite d'intervalles L_i, telle que $\Omega_1 \subset \bigcup_i L_i$.

Considérons, par exemple, une mesure μ définie sur les intervalles $\mathfrak{L}$ dans l'espace B des fonctions $x(t)$ continues sur (0,1). Soit (t_i) une suite de points de (0,1) convergeant vers τ. Considérons les intervalles

$$\mathfrak{L}_i : |x(t_i)| < i \quad , \quad i = 1, 2, \ldots$$

Il est clair que l'on peut supposer μ définie de telle sorte que $\mu(\mathfrak{L}_i) = \alpha_i$ avec $\Sigma\,\alpha_i < 1$ et que les conditions de compatibilité de Kolmogoroff soient satisfaites (μ pourra même être telle que les variables $x(t_i)$ soient indépendantes). Alors, on aura $\Omega_1 \subset \bigcup_i L_i$ puisque, $x(t)$ étant une fonction quelconque de B , elle a une valeur finie en τ et, comme elle est continue en τ, on aura pour i assez grand: $|x(t_i)| < i$. Mais on a cependant

$$m(\bigcup_{i=1}^{n} L_i) = \mu(\bigcup_{i=1}^{n} \mathfrak{L}_i) \leqslant \sum_{i=1}^{n} \alpha_i \leqslant \Sigma\,\alpha_i < 1$$

et par suite

$$m(\bigcup_i L_i) < 1$$

Ω_1 n'est donc pas de mesure extérieure égale à 1 et la mesure μ ne peut être étendue à K.

Le théorème limite central. Supposons maintenant que les variables indépendantes x_n aient toutes pour valeur moyenne θ (zéro de l'espace B), ou, ce qui revient au même, que $Mf(x_n) = 0$

Rappelons qu'on dira que $A^{(1)}$ est un ensemble de continuité pour $g^{(1)}$, si[2] pour tout $\varepsilon > 0$ donné, on peut trouver un entourage $V^{(1)}$ de Ω_1 tel que:

$$Q \in K^{(1)} \text{ \underline{et} } Q \subset V^{(1)}(A^{(1)}) \text{ \underline{implique} } g^{(1)}(Q) < g^{(1)}(A^{(1)}) + \varepsilon$$

(5)

$$R \in K^{(1)} \text{ \underline{et} } R \subset V^{(1)}(A'^{(1)}) \text{ \underline{implique} } g^{(1)}(R) < g^{(1)}(A'^{(1)}) + \varepsilon$$

où $A'^{(1)}$ désigne le complémentaire de $A^{(1)}$ par rapport à Ω_1 .

Soit donc $A^{(1)}$ un ensemble de continuité et $V^{(1)}$ l'entourage correspondant à ε au sens de (5). On a

$$V^{(1)}(A^{(1)}) = V(A^{(1)}) \cap \Omega_1 , \quad V(A^{(1)}) \supset \overline{A^{(1)}}$$

où $\overline{A^{(1)}}$ est la fermeture de $A^{(1)}$ dans Ω . Or $\overline{A^{(1)}}$, ensemble fermé de l'espace compact Ω , est compact et $V(A^{(1)})$ est ouvert; on a donc

$$\overline{A^{(1)}} \subset V(A^{(1)}) = \bigcup_{a \in V(A^{(1)})} L_a \tag{6}$$

où L_a désigne un intervalle voisinage du point a . Or les intervalles, comme on l'a déjà observé, sont ouverts et l'on tire de (6) qu'il existe des points $a_1, a_2, \ldots, a_k$, <u>en nombre fini</u>, tels que

$$\overline{A^{(1)}} \subset \bigcup_{i=1}^{k} L_{a_i} \subset V(A^{(1)}).$$

On aura

$$A^{(1)} \subset \bigcup_{i=1}^{k} L_{a_i} \cap \Omega_1 = \bigcup_{i=1}^{k} L_{a_i}^{(1)} \subset V^{(1)}(A^{(1)})$$

et par suite, en vertu de (5),

$$g^{(1)}\Big(\bigcup_{i=1}^{k} L_{a_i}^{(1)}\Big) < g^{(1)}(A^{(1)}) + \varepsilon . \tag{7}$$

De plus, $m_n^{(1)}$ convergeant vers $g^{(1)}$ sur toute réunion finie d'intervalles, on aura pour n assez grand,

$$m_n^{(1)}(A^{(1)}) \leqslant m_n^{(1)}\Big(\bigcup_{i=1}^{k} L_{a_i}^{(1)}\Big) < g^{(1)}\Big(\bigcup_{i=1}^{k} L_{a_i}^{(1)}\Big) + \varepsilon . \tag{8}$$

On tire de (7) et (8)

$$m_n^{(1)}(A^{(1)}) < g^{(1)}(A^{(1)}) + 2\varepsilon \tag{9}$$

et de même

$$m_n^{(1)}(A'^{(1)}) < g^{(1)}(A'^{(1)}) + 2\varepsilon \tag{10}$$

(2) Ceci est une adaptation des définitions données dans [V] au cas où les $V^{(1)}(A^{(1)})$ ne sont pas nécessairement mesurables.

Mais (10) s'écrit aussi

(11) $$g^{(1)} (A^{(1)}) < m_n^{(1)} (A^{(1)}) + 2\varepsilon$$

De (9) et (11) on déduit que l'on a, pour n assez grand,

$$| m_n^{(1)} (A^{(1)}) - g^{(1)} (A^{(1)}) | < 2\varepsilon$$

Ceci achève la démonstration.

BIBLIOGRAPHIE

[I] Bourbaki; Les structures fondamentales de l'analyse, livre III, Topologie générale, Chapitres II et III . Paris, Hermann, 1940.

[II] Cramer, H. ; Random Variables and Probability Distributions, Cambridge, 1937.

[III] Doob, J.L. and Ambrose, W. ; On twe formulations of the theory of stochastic processes depending upon a continuous parameter, Ann. of Math. vol. 41 (1940), p. 737-745.

[IV] Doob, J.L.. Probability in Function Space, Bull. Am. Math. Sec. vol. 53, Number 1, Part 1 (1947) p. 15-30.

[V] Doss, S. ; On stochastic Convergence in Uniform Spaces (à paraître)

[VI] Doss, S. ; Sur une estimation exhaustive pour la moyenne d'une variable aléatoire obéissant à la loi de Laplace dans un espace de Banach (voir dans ce fascicule).

[VII] Fréchet, M. ; Généralisations de la loi de probabilité de Laplace, Ann. Inst. H. Poincaré, t. XII, fasc.I (1951) p. 1-29.

[VIII] Kolmogoroff, A. ; Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin, Springer, 1933.

[IX] E. Mourier ; Eléments aléatoires laplaciens dans un espace de Banach; C.R. Acad. Sci. Paris T. 236 (1953) pp.375-376.

[X] E. Mourier et R. Fortet ; Lois des grands nombres pour des éléments aléatoires prenant leurs valeurs dans un espace de Banach; C.R. Acad. Sci. Paris, T. 237 (1953)pp. 18-20.

SUR LA CONVERGENCE PRESQUE COMPLÈTE DES MOYENNES DE VARIABLES ALÉATOIRES

(Théorèmes de HSU, ROBBINS et ERDÖS)

par

D. DUGUÉ

Dans un article des Proceedings of National Academy of Sciences 1947 p. 25 - 31 , Hsu et Robbins avaient établi la convergence presque complète d'un échantillon de variables aléatoires ayant F(x) pour loi de probabilité totale vers $\int x \, dF(x)$ si $\int x^2 dF(x)$ existe. Erdös avait ultérieurement démontré cette convergence d'une manière plus simple et établi que les conditions énoncées étaient nécessaires (Annals of Mathematical Statistics - Vol XX p. 286). J'ai moi-même établi que ces conditions étaient nécessaires par une autre méthode fondée sur la croissance de la plus grande valeur d'un échantillon (publiée dans les Actes du 28ème Congrès de l'Institut International de Statistique qui doivent prochainement paraître). Je vais indiquer ici une démonstration du fait que ces conditions sont suffisantes et qui me paraît plus simple que les deux démonstrations données précédemment. Cette démonstration s'inspire de la démonstration de Kolmogoroff pour la convergence presque certaine.

Rappelons qu'une suite de variables X_1, X_2, ... X_n.... converge presque complètement vers 0 (Hsu et Robbins qui ont introduit cette notion lui donne le nom de convergence complète, faute d'un meilleur terme, disaient-ils : je préfère pour ma part parler de convergence presque complète), si la série de terme général $1 - F_n(x) + F_n(-x)$ est convergente quel que soit $x > 0$ avec $F_n(x) = \text{Prob}\,(X_n < x)$.

Dans le cas où les variables X_1, X_2, ... X_n sont indépendantes cette condition est équivalente à la convergence presque certaine. Mais d'une manière générale elle n'est pas nécessaire pour assurer la convergence presque certaine. La condition énoncée pour la convergence <u>presque</u> complète entraîne que toute suite de variables Y_1, Y_2, ... Y_n, ... <u>indépendantes ou non</u>, telles que Y_n ait même loi de probabilité que X_n (dans le mémoire cité je propose de dire que dans ce cas X_n et Y_n sont légalement superposables) converge presque certainement vers 0.

Il s'agit donc de démontrer que sous les conditions que $\int x \, d \, F(x)$ et $\int x^2 d \, F(x)$ existent $\frac{X_1 + \cdots + X_n}{n}$ converge presque complètement vers $\int x \, d \, F(x)$, X_i ayant F(x) pour loi de probabilité totale.

Cela revient à montrer que la suite de variables aléatoires

$$\frac{X_1^{(n)} + X_2^{(n)} + \cdots + X_n^{(n)}}{n}$$

converge presque complètement vers $\int x \, d\, F(x)$, $X_i^{(n)}$ et $X_j^{(p)}$ étant indépendants, quels que soient n, p, i et j.(les variables étant indépendantes la convergence presque complète est identique dans ce cas à la convergence presque certaine).

Prenons une suite de variables $Y_1^{(1)}$, $Y_1^{(2)}$, $Y_2^{(2)}$, ...$Y_1^{(n)}$ ... $Y_n^{(n)}$ ... indépendantes et ainsi définies :

$$Y_i^{(n)} = X_i^{(n)} \text{ si } |X_i^{(n)}| < n \text{ et } Y_i^{(n)} = 0 \text{ si } |X_i^{(n)}| > n$$

A) Montrons tout d'abord que

$$S_n = \frac{Y_1^{(n)} + \cdots + Y_n^{(n)}}{n}$$ converge presque complètement,

sous les conditions indiquées, vers $\int x \, d\, F(x)$.

Ce fait est une conséquence des résultats de Borel et de Cantelli utilisant l'inégalité de Tchebycheff et par conséquent pouvant être interprétés comme des démonstrations de la convergence presque complète.

Si la série des moments du 4° ordre de $S_n - E(S_n)$ converge et si $E(S_n)$ tend vers une limite a , la série ayant pour terme général

$\text{Prob}\{|S_n - a| > \varepsilon\}$ converge quel que soit ε positif

$E(S_n)$ tend vers $\int x dF(x)$ car :

$E(S_n) = \frac{1}{n}\left[E(Y_1^{(n)}) + \cdots + E(Y_n^{(n)})\right]$ et $E(Y_i^{(n)}) = \int_{-n}^{+n} x dF(x)$ tend vers $\int x \, d\, F(x)$ quand n augmente indéfiniment. Il en est donc de même de $E(S_n)$.

De plus pour une même valeur de n tous les $Y_i^{(n)}$ ayant même loi de probabilité on a :

$$E(S_n - E(S_n))^4 = \frac{1}{n^3} E(Y_i^{(n)} - E\, Y_i^{(n)})^4 + \frac{3(n-1)}{n^3} E\left[(Y_i^{(n)} - E(Y_i^{(n)}))^2\right]^2$$

Posons : $u_n = \frac{1}{n^3} E(Y_i^{(n)} - E(Y_i^{(n)}))^4$ et $v_n = \frac{3(n-1)}{n^3}\left[E(Y_i^{(n)} - E(Y_i^{(n)})^2\right]^2$

Il s'agit de montrer que $\sum u_n$ et $\sum v_n$ sont convergents.

La série $\sum v_n$ converge puisque $E(Y_i^{(n)} - E(Y_i^{(n)}))^2 < \int x^2 d\, F(x)$

On a :

$$u_n = E\left[Y_i^{(n)} - E(Y_i^{(n)})\right]^4 = \int_{-n}^{+n} x^4 dF(x) - 4E(Y_i^{(n)}) \int_{-n}^{+n} x^3 dF(x)$$
$$+ 6\left[E(Y_i^{(n)})\right]^2 \int_{-n}^{+n} x^2 dF(x)$$
$$- 4\left[E(Y_i^{(n)})\right]^3 \int_{-n}^{+n} x\, dF(x) + \left[E(Y_i^{(n)})\right]^4$$

Comme $\int_{-n}^{+n} x\, dF(x)$ et $\int_{-n}^{+n} x^2 dF(x)$ tendent vers une limite quand n augmente indéfiniment, la démonstration de la convergence de la série $\sum u_n$ revient à établir que :

$$\sum \frac{1}{n^3} \int_{-n}^{+n} x^4 dF(x) \text{ et } \sum \frac{1}{n^3} \int_{-n}^{+n} x^3 dF(x)$$

sont des séries convergentes

Si $\int x^2 dF(x)$ existe, $\left[\int_{-n}^{+n} x^3 dF(x)\right] < n \int_{-n}^{+n} x^2 dF(x)$ donc la secondesérie converge.

On a :

$$\sum \frac{1}{n^3} \int_{-n}^{+n} x^4 dF = \sum_{p=1}^{\infty} \left\{ \left[\sum_{n=p}^{\infty} \frac{1}{n^3}\right] \int_{p-1}^{p} x^4 dF(x) \right\}$$
$$+ \sum_{p=1}^{\infty} \left\{ \left[\sum_{n=p}^{\infty} \frac{1}{n^3}\right] \int_{-p}^{-(p-1)} x^4 dF(x) \right\}$$

On a : $\sum_{n=p}^{\infty} \frac{1}{n^3} < \frac{2}{p^2}$ et par suite :

$$\sum \frac{1}{n^3} \int_{-n}^{+n} x^4 dF(x) < 2 \sum_{p=1}^{\infty} \int_{p-1}^{p} x^2 dF(x) + 2 \sum_{p=1}^{\infty} \int_{-p}^{-(p-1)} x^2 dF(x) = 2 \int_{-\infty}^{+\infty} x^2 dF(x)$$

La série dont le terme général est Prob $\left[\left| S_n - \int x\, dF(x) \right| > \varepsilon\right]$ est donc convergente.

B) Il suffit maintenant de montrer que sous les hypothèses faites

$$\frac{X_1^{(n)} + \cdots\cdots + X_n^{(n)}}{n} - S_n \quad \text{converge presque complètement vers 0.}$$

Or :

$$\frac{X_1^{(n)} + \cdots\cdots + X_n^{(n)}}{n} - S_n = \frac{X_1^{(n)} - Y_1^{(n)} + \cdots\cdots + X_n^{(n)} - Y_n^{(n)}}{n}$$

Si quel que soit i on a $\left| X_i^{(n)} - Y_i^{(n)} \right| < x$ on a :

$$\left| \frac{X_1^{(n)} + \cdots\cdots + X_n^{(n)}}{n} - S_n \right| < x$$

Donc puisque $X_i^{(n)} - Y_i^{(n)}$ est indépendant de $X_j^{(n)} - Y_j^{(n)}$

$$\text{Prob}\left[\left\{\left|X_i^{(n)} - Y_i^{(n)}\right| < x\right\}\right]^n < \text{Prob}\left\{\left|\frac{X_1^{(n)} + \cdots + X_n^{(n)}}{n} - S_n\right| < x\right\}$$

$$\text{et Prob}\left\{\left|\frac{X_1^{(n)} + \cdots + X_n^{(n)}}{n} - S_n\right| \geqslant x\right\} < 1 - \left[\text{Prob}\left\{\left|X_i^{(n)} - Y_i^{(n)}\right| < x\right\}\right]^n$$

$$< n\left[1 - \left[\text{Prob}\left\{\left|X_i^{(n)} - Y_i^{(n)}\right| < x\right\}\right]\right]$$

$$\text{Prob}\left\{\left|X_i^{(n)} - Y_i^{(n)}\right| < x\right\} > F(n) - F(-n)$$

et finalement :

$$\text{Prob}\left\{\left|\frac{X_1^{(n)} + \cdots + X_n^{(n)}}{n} - S_n\right| \geqslant x\right\} < n\left[1 - F(n) + F(-n)\right]$$

Si $\int x^2 dF(x)$ existe on sait que la série de terme général

$$n\left[1 - F(n) + F(-n)\right]$$

converge ce qui établit la convergence presque complète de $\frac{X_1^{(n)} + \cdots + X_n^{(n)}}{n} - S_n$ vers 0 et par suite le théorème.

LES PROCESSUS STOCHASTIQUES [1] EN DÉMOGRAPHIE

par

D. D. JOSHI

INTRODUCTION

La théorie classique de la démographie emploie, pour résoudre les divers problèmes posés, une méthode de raisonnement que nous appellerons ici la méthode déterministe. Cette méthode se base sur l'hypothèse suivante :

> "L'état" de la population à n'importe quel instant t peut être prévu si l'on connaît son "état" à un instant antérieur $t_0 < t$; les seules incertitudes dans cette prévision sont dues à la variabilité des taux de mortalité et reproduction.

La théorie a été développée par plusieurs auteurs et une étude systématique et détaillée se trouve dans Lotka (10), Kostitzin [9] , et Rhodes [12](2).

Cette méthode a trouvé pas mal d'applications et a bien servi à résoudre plusieurs problèmes pratiques. Mais il est évident qu'une étude plus satisfaisante du problème doit tenir compte des variations dues au hasard et ici la théorie déterministe ne nous sert pas. C'est seulement une méthode probabiliste qui pourra nous fournir une idée de la grandeur de ces variations.

Le point de départ de la méthode probabiliste est le suivant:

> Etant donné "l'état" de la population à l'instant t_0, il existe une certaine probabilité pour que la population se trouve dans un "état" déterminé quelconque à un instant $t > t_0$.

Ainsi, au lieu de considérer l'état dans lequel se trouvera la population à un instant postérieur t, on considère la 'valeur moyenne' des états possibles à l'instant t et les variations aléatoires autour de cette valeur moyenne. Vue ainsi, l'évolution d'une population humaine peut être considérée comme un processus stochastique.

---

(1) Cet article est le développement d'un mémoire présenté à l'Institut de Statistique de l'Université de Paris en Juin 1953. Nous exprimons ici nos reconnaissances respectueuses à Monsieur le Professeur Georges Darmois à qui nous devons l'inspiration de cette étude et qui nous a constamment guidé avec ses conseils.

(2) Les chiffres entre crochets renvoient à la bibliographie à la fin.

Depuis peu, divers auteurs ont utilisé la théorie des processus stochastiques pour étudier des problèmes semblables qui se posent en biologie. On en trouvera une bibliographie très complète dans Kendall [8] . Cependant, quant à la démographie, jusqu'à présent on ne s'est occupé que du problème de croissance d'une population pour lequel divers modèles stochastiques ont été proposés. A notre connaissance ce travail, bien qu'élémentaire, est la première étude systématique des problèmes de la démographie mathématique du point de vue de la théorie des processus stochastiques.

La première partie de cet article donne un résumé des résultats bien connus de la théorie classique. Dans la seconde partie après avoir développé le modèle du processus simple de naissance et de mort selon la méthode employée par Kendall [6] , nous considérons le problème d'estimation des paramètres qui définissent le processus. Dans le cas général, la solution de ce problème a été obtenue par Kendall [8] et Moran [11] . Nous avons montré qu'il existe une solution simple pour le cas démographique, qu'on peut obtenir par la méthode du maximum de vraisemblance. Dans la troisième partie nous avons étendu ce modèle au cas d'une population comprenant deux types d'individus. Nous avons montré qu'en utilisant un modèle stochastique à deux variables on peut étudier la différence entre le nombre d'individus du sexe masculin et celui du sexe féminin dans une population humaine. Le problème d'estimation dans ce cas a été résolu par la méthode déjà utilisée dans la seconde partie. Enfin, dans la quatrième partie, nous généralisons le problème au cas où les taux de natalité et de mortalité dépendent de l'âge de l'individu. Bien que ce problème soit encore loin d'être résolu complètement quelques résultats intéressants ont été obtenus, en utilisant la théorie des processus de Poisson généralisés.

PREMIÈRE PARTIE

LA THÉORIE DÉTERMINISTE

La théorie déterministe, due en grande partie à Lotka, se base sur les fonctions suivantes:

$N(t)$ le nombre de naissances "annuelles" à l'époque t,
$M(t)$ le nombre des morts "annuelles" à l'époque t,
$n(t)$ le taux de natalité,
$m(t)$ le taux de mortalité,
$p(a)$ la fonction de survie, qui est la proportion des individus nés à l'époque $t-a$ qui sont encore en vie à l'époque t, ayant alors l'âge a. Cette fonction qui, en général, dépend de l'époque t, est supposée indépendante du temps,
$f(a)$ la fécondité effective du sexe féminin à l'âge a, définie comme le nombre de filles nées vivantes par an "per capita" de femmes d'âge a. On la suppose aussi indépendante du temps. On introduit aussi les fonctions $f_1(a)$, $f_2(a)$,..... ...$f_n(a)$ où $f_n(a)$ est le nombre de n-ièmes naissances par femme d'âge a.

A partir de ces fonctions on obtient des résultats bien connus de la démographie mathématique. Ainsi, si l'on suppose que les taux de natalité et de mortalité sont constants et ne dépendent pas du temps on a la loi exponentielle de la croissance ou loi malthusienne. Si, au contraire, on suppose que $n(t)$ et $m(t)$ sont fonctions linéaires du nombre d'individus à l'instant t,on a la loi logistique de croissance.

Si l'on désigne par $\gamma(t)$ le nombre d'individus que comprend la population à l'époque t, on a les relations suivantes:

$$\frac{d\gamma}{dt} = N(t) - M(t), \tag{1}$$

$$\gamma(t) = \int_0^\infty N(t-a)p(a)da, \tag{2}$$

$$N(t) = \int_0^\infty N(t-a)p(a)f(a)da, \tag{3}$$

$$M(t) = -\int_0^\infty N(t-a)\frac{d}{da}p(a)da \tag{4}$$

On définit encore un coefficient de répartition par âge $c(a,t)$, tel que le nombre d'individus compris entre les âges a et a+da soit $\gamma(t)c(a,t)da$. Il s'ensuit que

$$\gamma(t)c(a,t)da = N(t-a)p(a)da, \tag{5}$$

(6) $$c(a,t) = \frac{N(t-a)}{\gamma(t)}\, p(a),$$

et que

(7) $$\int_0^\infty c(a,t)da = 1.$$

On a aussi les relations suivantes qui lient les taux de natalité et de mortalité au coefficient de répartition par âge.

(8) $$n(t) = \frac{N(t)}{\gamma(t)} = c(0,t)$$

(9) $$m(t) = -\int_0^\infty c(a)\frac{d\log p(a)}{da}\, da$$

Si le coefficient c(a,t) est indépendant de t, tel que l'on peut l'écrire simplement c(a), on a

(10) $$n(t) = c(0) = \text{constante}$$

(11) $$m(t) = -\int_0^\infty c(a)\frac{d\log p(a)}{da} da = \text{constante}.$$

Et ainsi

(12) $$\frac{1}{\gamma}\,\frac{d\gamma}{dt} = n(t) - m(t) = \text{constante} = r$$

ou

(13) $$\gamma = \gamma_0 e^{rt}$$

et la population croît (ou décroît) d'après la loi de Malthus. Si, au lieu de supposer que le coefficient c(a,t) ne dépend pas de t, on avait supposé que les taux de natalité et de mortalité sont constants, il se serait ensuit que le coefficient c(a,t) ne dépend pas de t. Ainsi donc une population malthusienne est caractérisée ou par l'indépendance par rapport au temps des taux de natalité et de mortalité ou par l'indépendance par rapport au temps du coefficient de répartition par âge.

Pour une population malthusienne on a, de plus

(14) $$N(t) = n.\gamma(t) = n\gamma_0 e^{rt} = N_0 e^{rt}$$

(15) $$M(t) = m\gamma(t) = m\gamma_0 e^{rt} = M_0 e^{rt}$$

tel que les naissances et les décès annuels obéissent à la même loi.

Etudions maintenant l'équation importante

(16) $$N(t) = \int_0^\infty N(t-a)p(a)f(a)da$$

dite équation de Lotka. On sait bien qu'il existe une solution sous la forme de la série exponentielle

(17) $$N(t)=Q_1e^{r_1t}+Q_2e^{r_2t}+Q_3e^{r_3t}+\cdots\cdots$$

où les coefficients r_n sont les racines de l'équation

(18) $$1=\int_0^\infty e^{-ra}\,p(a)f(a)da$$

et les coefficients Q_n dépendent des conditions initiales.

La fonction p(a)f(a) étant réelle et positive, on démontre que l'équation (18) ne peut avoir qu'une seule racine réelle ρ, telle que

$$\rho \gtrless 0 \text{ suivant que } R_0 \gtrless 1$$

où

(19) $$R_0=\int_0^\infty p(a)f(a)da.$$

Il est évident que R_0, que l'on appelle le taux net de reproduction, est le nombre moyen d'enfants qu'a une femme pendant la durée de sa vie.
On démontre aussi que pour toute racine complexe

$$r = u + iv$$

on a

$$u<\rho$$

de sorte que l'amplitude relative des oscillations, par rapport au terme apériodique $Q_\rho e^{\rho t}$ va en diminuant, et, en fin de compte c'est ce terme apériodique qui domine l'évolution des naissances car par rapport à lui tous les autres termes de la solution deviennent négligeables. On a alors

(20) $$N(t) = Q_\rho e^{\rho t}$$

Du moment que la formule (20) est applicable, on a

(21) $$\gamma\,(t)=\int_0^\infty N(t-a)p(a)da$$

$$= ke^{\rho t} \qquad (k=\text{une constante}).$$

Nous avons donc, dès lors, une population malthusienne augmentant ou diminuant selon le taux d'accroissement constant ρ. Pourvu que la mortalité et la fécondité soient indépendantes du temps, le taux d'accroissement s'approche de plus en plus de la valeur asymptotique ρ, ou taux intrinsèque d'accroissement, quelle que soit la répartition initiale par âge et les taux de natalité et de mortalité. Ainsi ρ exprime la capacité fondamentale de multiplication de la population, libérée de l'influence perturbatrice d'une répartition initiale arbitraire par âge.

Nous terminons ici ces remarques brèves et rudimentaires sur la théorie classique. Dans la seconde partie nous allons considérer ces problèmes du point de vue probabiliste.

DEUXIÈME PARTIE

LA THÉORIE STOCHASTIQUE

Le processus simple de naissance et de mort

La théorie des processus stochastiques, celle des processus markoviens en particulier, nous permet de construire des modèles semblables aux modèles déterministes, pour étudier les divers problèmes démographiques. Nous emploierons les méthodes déjà utilisées par divers auteurs pour é udier le développement des colonies de bactéries en les modifiant selon les conditions spécifiques des problèmes de la démographie humaine. Pour construire ces modèles nous n'avons besoin, pour le moment, que des deux fonctions, les taux de natalité et de mortalité, que nous définissons ci-dessous.

LES HYPOTHÈSES

Soit $\lambda(\alpha,t)$ et $\mu(\alpha,t)$ les taux par âge de natalité et de mortalité respectivement définis ainsi qu'il suit:

(i) un individu en vie à l'instant t et ayant l'âge α, a la probabilité

$$\lambda(\alpha,t)dt+o(dt)$$

de donner naissance à un autre individu (de l'âge 0) dans l'intervalle de temps $(t, t+dt)$;

(ii) un individu en vie à l'instant t et ayant l'âge α, à la probabilité

$$\mu(\alpha,t)dt+o(dt)$$

de mourir dans l'intervalle de temps $(t, t+dt)$.

De plus, nous supposons que les sous-populations produites par deux individus de la population se développent indépendamment l'une de l'autre et que la probabilité de naissance et la probabilité de mort sont indépendantes, la probabilité de reproduction et de mort simultanée (ainsi que la probabilité de plusieurs naissances) étant considérées comme inférieures à dt. Aussi nous ne nous occupons que de la partie féminine de la population et supposons que le nombre d'individus du sexe masculin s'ajuste en conséquence.

LE PROCESSUS SIMPLE DE NAISSANCE ET DE MORT

Supposons d'abord que les fonctions $\lambda(\alpha,t)$ et $\mu(\alpha,t)$ ne dépendent ni de l'époque t ni de l'âge α telles que l'on peut les considérer comme les constantes λ et μ. Nous avons en ce cas le processus simple de naissance et de mort. (voir Kendall [6]).

Soit γ_t la variable aléatoire qui désigne le nombre d'individus dans la population à l'instant t et soit $P_n(t)$ la probabilité pour que $\gamma_t = n$, et supposons que $\gamma_o = n_o$.

La probabilité $P_n(t)$ satisfait aux equations différentielles

(22) $$\frac{\partial P_n(t)}{\partial t} = -n(\lambda+\mu)P_n(t)+(n-1)\lambda P_{n-1}(t)+(n+1)\mu P_{n+1}(t), \quad n \geqslant 1$$

(23) $$\frac{\partial P_o(t)}{\partial t} = \mu P_1(t),$$

avec les conditions initiales

(24) $$P_{n_o}(0)=1 \; ; \quad P_n(0)=0, \quad n \neq n_o.$$

En prenant $P_n(t)=0$ pour $n<0$, la fonction génératrice

(25) $$\varphi(z,t) = \sum_{-\infty}^{+\infty} P_n(t)\, z^n$$

satisfait à l'équation linéaire aux dérivées partielles

(26) $$\frac{\partial \varphi}{\partial t} = (\lambda z - \mu)(z-1)\frac{\partial \varphi}{\partial z}$$

avec

(27) $$\varphi(z,0) = z^{n_o} \; ; \; \varphi(1,t) = 1.$$

La solution générale de (26) est de la forme

(28) $$\varphi(z,t) = \Phi\left\{\frac{\lambda z-\mu}{z-1} \; e^{-(\lambda-\mu)t}\right\}$$

et tenant compte des conditions initiales, on a

(29) $$\varphi(z,t) = \left[\frac{(\lambda z-\mu)+(1-z)\mu\, e^{(\lambda-\mu)t}}{(\lambda z-\mu)+(1-z)\lambda\, e^{(\lambda-\mu)t}}\right]^{n_o}$$

D'où on trouve

$$P_o(t) = \left[\frac{\mu e^{(\lambda-\mu)t}-\mu}{\lambda e^{(\lambda-\mu)t}-\mu}\right]^{n_o} = \xi^{n_o}$$

(30) $$P_r(t) = P_o(t)\,\eta^r\left[\sum_{x=o}^{r}\zeta^x\binom{n_o}{x}\frac{\Gamma(n_o+r-x)}{\Gamma(r-x+1)\Gamma(n_o)}\right] \qquad r<n_o$$

$$P_r(t) = P_o(t)\,\eta^r\left[\sum_{x=o}^{r}\zeta^x\binom{n_o}{x}\frac{\Gamma(n_o+r-x)}{\Gamma(r-x+1)\Gamma(n_o)}\right] \qquad r \geqslant n_o$$

où ξ, η, ζ sont donnés par les relations suivantes:

(31) $$\frac{\xi}{\mu} = \frac{\eta}{\lambda} = \frac{e^{(\lambda-\mu)t}-1}{\lambda e^{(\lambda-\mu)t}-\mu}, \quad \zeta = \frac{1-\xi-\eta}{\xi\eta}.$$

Si la valeur initiale $n_0=1$, on a

(32) $$P_o(t) = \xi$$
$$P_n(t) = (1-\xi)(1-\eta)\eta^{n-1} \qquad n \geqslant 1$$

De (29) on a

(33) $$\overline{\gamma}_t = E\left[\gamma_t\right] = \frac{\partial}{\partial z}\varphi(1,t) = n_o e^{(\lambda-\mu)t}$$

(34) $$\text{Var}\ \left[\gamma_t\right] = \frac{\partial^2}{\partial z^2}\varphi(1,t) - \overline{\gamma}_t\left(\overline{\gamma}_t - 1\right) = n_o\ \frac{\lambda+\mu}{\lambda-\mu}\ e^{(\lambda-\mu)t}\ \left\{ e^{(\lambda-\mu)t} - 1 \right\}$$

Ainsi l'espérance mathématique du nombre d'individus suit la loi exponentielle et nous retrouvons la loi malthusienne.

PROBABILITÉ D'EXTINCTION DE LA POPULATION

Examinons maintenant les expressions pour l'espérance mathématique et la variance de γ_t. L'espérance mathématique nous donne le nombre moyen d'individus que comprend la population à l'instant t. On voit que ce nombre moyen croît exponentiellement vers l'infini si $\lambda>\mu$, décroît exponentiellement vers zéro si $\lambda<\mu$ et est constante si $\lambda = \mu$. La variance de γ_t tend vers l'infini si $\lambda \geqslant \mu$, et tend vers zéro si $\lambda<\mu$, lorsque t tend vers l'infini.

Considérons ensuite la limite de φ (z,t) lorsque t tend vers l'infini. Nous avons

(35) $$\lim_{t\to\infty} \varphi(z,t) = 1 \qquad \lambda \leqslant \mu$$
$$= \left[\frac{\mu}{\lambda}\right]^{n_0} \qquad \lambda > \mu$$

Cela montre que

(36) $$\lim_{t\to\infty} P_0(t) = 1 \qquad \lambda \leqslant \mu$$
$$= \left[\frac{\mu}{\lambda}\right]^{n_0} \qquad \lambda > \mu$$

et que

(37) $$\lim_{t\to\infty} P_n(t) = 0 \qquad n>0 \text{ mais fini}$$

Ces limites montrent que pour $\lambda \leqslant \mu$ l'extinction de la population est "presque certaine"et que pour $\lambda>\mu$ il y a une probabilité non nulle $(=\left[\frac{\mu}{\lambda}\right]^{n_o})$ pour que la population s'éteigne. De plus, dans ce dernier cas ($\lambda>\mu$) la population augmente sans cesse avec la probabilité $1 - (\mu/\lambda)^{n_o}$.

Ainsi nous voyons que pour étudier l'extinction d'une population il ne suffit pas de considérer la valeur asymptotique du nombre moyen des individus. Pour la théorie déterministe une population croît sans cesse vers l'infini si le taux d'accroissement est positif, et décroît vers zéro si le taux d'accroissement est négatif. La méthode probabiliste, au contraire, nous donne une idée plus précise de ce phénomène et nous voyons que même dans le cas où $\lambda = \mu$ (c'est-à-dire $E(\gamma_t)$ = constante) l'extinction de la population au cours du temps est presque certaine

RÉPARTITION PAR AGE

Désignons par $d\gamma(\alpha,t)$ le nombre d'individus dont l'âge à l'instant t est compris entre les âges α et $\alpha + d\alpha$, tel que l'intégrale de Stieltjes par rapport à α

$$\int_{\alpha_1}^{\alpha_2} d\gamma(\alpha,t)$$

est le nombre d'individus dans le groupe d'âges (α_1,α_2). On suppose que la répartition initiale par âge $\gamma(\alpha,0)$ est connue. Si pour les k intervalles non-empiétants d'âges (α_r,β_r), r=1,2..,k on peut trouver la loi des k variables aléatoires

$$\int_{\alpha_r}^{\beta_r} d\gamma(\alpha,t) \qquad (r=1,2,\ldots\ldots,k)$$

le problème est résolu.

Nous ne nous occuperons pas ici de ce problème qui a été traité en détail par Kendall [6] et [7] . Cependant, l'espérance mathématique et la variance de $d\gamma(\alpha,t)$ peuvent être obtenues par un raisonnement simple.

Considérons premièrement le cas où $\alpha < t$. Nous supposons que

$$\begin{aligned} d\gamma(\alpha,t) &= 0 \quad \text{avec la probabilité } 1 - \delta(\alpha,t)\,d\alpha + O(d\alpha) \\ &= 1 \quad \text{avec la probabilité } \delta(\alpha,t)\,d\alpha + O(d\alpha) \\ &\geqslant 2 \quad \text{avec la probabilité } O(d\alpha) \end{aligned}$$

Alors, au premier ordre, $d\gamma(\alpha,t)$ suit la loi de Poisson dont l'espérance mathématique et la variance sont approximativement égales à $\delta(\alpha,t)d\alpha$.

Le nombre d'individus $d\gamma(\alpha,t)$ dépend du nombre des survivants d'individus nés dans l'intervalle de temps $(t-\alpha, t-\alpha+d\alpha)$.

L'espérance mathématique du nombre d'individus que comprend la population à l'instant $(t-\alpha)$ est donnée par (33)

$$n_0 e^{(\lambda-\mu)(t-\alpha)}$$

n_0 étant le nombre initial d'individus. Donc l'espérance mathématique du nombre d'individus nés dans l'intervalle de temps $(t-\alpha,\ t-\alpha+d\alpha)$ est

$$n_0 e^{(\lambda-\mu)(t-\alpha)} \lambda d\alpha \, .$$

La probabilité pour qu'un individu né à l'instant $(t-\alpha)$ soit encore en vie à l'instant t est d'après le résultat (42) obtenu plus tard

$$e^{-\alpha\mu}$$

Donc

(38) $$E\left[d\gamma(\alpha,t)\right]=n_0\lambda e^{(\lambda-\mu)(t-\alpha)}e^{-\mu\alpha}\,d\alpha$$
$$=\mathrm{Var}\left[d\gamma(\alpha,t)\right]$$

C'est-à-dire

(39) $$E\left[d\gamma(\alpha,t)\right]=\bar{\gamma}_t\,\lambda e^{-\lambda\alpha}\,d\alpha$$

Ainsi le rapport de l'espérance mathématique du nombre d'individus appartenant au groupe d'âges $(\alpha,\alpha+d\alpha)$ à l'espérance mathématique du nombre total de la population ne dépend pas de t, et nous retrouvons la propriété qui caractérise une population malthusienne.

Dans le cas où $\alpha > t$, $d\gamma(\alpha,t)$ ne dépend que du nombre $d\gamma(\alpha-t,0)$ d'individus dont l'âge à l'instant t=0 est compris entre $\alpha-t$ et $\alpha-t+d\alpha$, les naissances pendant l'intervalle $(0,t)$ n'y intervenant pas. Donc $d\gamma(\alpha,t)$ suit maintenant la loi binome avec

$$"n" = d\gamma(\alpha-t,0)\ ,\quad "p"=e^{-\mu t}$$

et

(40) $$E\left[d\gamma(\alpha,t)\right] = e^{-\mu t}\,d\gamma(\alpha-t,0)$$

(41) $$\mathrm{Var}\left[d\gamma(\alpha,t)\right] = e^{-\mu t}\left(1-e^{-\mu t}\right)d\gamma(\alpha-t,0).$$

LA DURÉE DE VIE DE L'INDIVIDU

Cherchons d'abord la probabilité p (t) pour qu'un individu né à l'instant t=0 soit encore en vie à l'instant t. La probabilité conditionnelle pour qu'un individu existant à l'instant t soit encore en vie à l'instant t+dt est

$$1-\mu\,dt+0(dt)$$

Donc

$$p(t+dt)=p(t)\left[1-\mu dt+0(dt)\right]$$

$$\frac{dp(t)}{dt} = -\mu\,p(t)$$

avec la condition initiale

$$p(0) = 1$$

Donc

(42) $$p(t) = e^{-\mu t}$$

La probabilité p(t) correspond à la fonction de survie p(a) utilisée déjà dans la première partie.

La probabilité pour que la durée de vie d'un individu soit égale à T est alors la probabilité pour qu'un individu né à l'instant τ quelconque soit encore en vie à l'instant $\tau + T$ et soit mort pendant l'intervalle dT qui suit. Ainsi la probabilité pour que la durée de vie soit égale à T est donnée par

(43) $$\mu e^{-\mu T} dT \qquad 0 < T < \infty$$

LA LOI DE PROBABILITÉ DU NOMBRE DE DESCENDANTS

Cherchons maintenant la probabilité p_r considérée au moment de sa naissance (les morts-nés étant exclus) pour qu'un individu ait eu r(=0,1,2,3,.....) descendants (ni moins ni plus) avant sa mort.

Considérons d'abord la probabilité pour qu'un individu ait r descendants pendant le temps (calculé de moment de sa naissance) t. Cette probabilité $\pi_r(t)$ satisfait aux équations différentielles

$$\frac{\partial \pi_r(t)}{\partial t} = -\lambda \pi_r(t) + \lambda \pi_{r-1}(t)$$

$$\frac{\partial \pi_o(t)}{\partial t} = -\lambda \pi_o(t)$$

avec la condition initiale

$$\pi_o(0)=1 \ ; \quad \pi_r(0)=0, \quad r > 0 \ .$$

Ces équations ont pour solution

(44) $$\pi_r(t) = \frac{1}{r!} (\lambda t)^r e^{-\lambda t}$$

La probabilité pour qu'un individu ait la durée de vie T, et ait eu pendant cette durée de vie r descendants (ni plus ni moins) est, alors

(45) $$\frac{1}{r!} \mu (\lambda T)^r e^{-(\lambda+\mu)T} dT \ .$$

Donc

$$p_r = \int_0^\infty \frac{1}{r!} \mu (\lambda T)^r e^{-(\lambda+\mu)T} dT \ .$$

C'est-à-dire

(46) $$p_r = \left(1 - \frac{\lambda}{\lambda+\mu}\right)\left(\frac{\lambda}{\lambda+\mu}\right)^r, \qquad r = 0,1,2,...$$

On constate que cette probabilité ne dépend que du rapport λ/μ.

La fonction génératrice f (x) des probabilités p_r est

$$f(x) = \sum_{r=0}^{\infty} p_r x^r = \frac{\mu}{\lambda(1-x)+\mu}$$

D'après les résultats bien connus (c.f. Steffensen [13])la probabilité d'extinction d'une population (quand le nombre initial $n_0=1$) est donnée par la racine positive la plus petite de l'équation

$$x = f(x)$$

c'est-à-dire

$$x = \frac{\mu}{\lambda(1-x) + \mu}$$

ou

$$(\lambda x - \mu)(x-1) = 0,$$

et cela vérifie les résultats déjà trouvés.

ESTIMATION DES PARAMÈTRES λ ET μ

Moran [11] et suivant sa méthode Kendall [8] ont obtenu respectivement l'estimation de la somme $(\lambda+\mu)$ et du rapport λ/μ. Ces deux estimations sont basées sur deux échantillons indépendants obtenus par deux méthodes différentes et on ne sait pas si les estimations de λ et de μ tirées des précédentes seront les plus efficaces et absolument correctes. Nous ne considérons pas ici le problème général d'estimation des paramètres associés à un processus stochastique, mais nous allons obtenir une solution exacte et simple dans le cas spécial d'une population humaine. Cette solution est obtenue par la méthode du maximum de vraisemblance en utilisant la loi (45) du nombre de descendants qu'a un individu pendant la durée de sa vie (compte tenu du fait que cette loi est tirée des mêmes hypothèses qui définissent le processus).

Considérons un échantillon de n individus ayant la durée de vie

$$T_1, T_2, \ldots\ldots\ldots\ldots, T_n$$

et supposons que chacun de ces individus ait eu pendant la durée de sa vie

$$r_1, r_2, \ldots\ldots\ldots\ldots, r_n$$

descendants respectivement.

D'après (45) la vraisemblance est donnée par

$$\mathcal{L} = \frac{\mu^n}{(r_1!)(r_2!)\cdots(r_n!)} (\lambda T_1)^{r_1}(\lambda T_2)^{r_2}\cdots\cdots(\lambda T_n)^{r_n}$$
$$\times\ e^{-(\lambda+\mu)(T_1+T_2+\cdots\cdots+T_n)}.$$

C'est-à-dire

$$\log \mathcal{L} = \text{constante} + R\log\lambda - (\lambda+\mu)T + n\log\mu\ ,$$

où

$$R = r_1 + r_2 + \ldots\ldots + r_n ,$$
$$T = T_1 + T_2 + \ldots\ldots + T_n .$$

Nous avons alors

$$\frac{\partial \log \mathcal{L}}{\partial \lambda} = \frac{R}{\lambda} - T , \quad \frac{\partial \log \mathcal{L}}{\partial \mu} = \frac{n}{\mu} - T .$$

Donc les estimations, des paramètres λ et μ sont données par

(47) $$\hat{\lambda} = \frac{R}{T} , \quad \hat{\mu} = \frac{n}{T} .$$

On démontre facilement que le couple $(\hat{\lambda} , \hat{\mu})$ suit asymptotiquement une loi de Laplace-Gauss avec les valeurs moyennes λ et μ respectivement, et la matrice des covariances

(48) $$\begin{Vmatrix} \frac{\lambda\mu}{n} & 0 \\ 0 & \frac{\mu^2}{n} \end{Vmatrix}$$

TROISIÈME PARTIE

LA THÉORIE STOCHASTIQUE

Population comprenant deux types d'individus

Jusqu'à présent nous ne nous étions occupés que de la partie féminine de la population et nous avions supposé que le nombre d'individus du sexe masculin s'ajuste en conséquence. C'est d'ailleurs la méthode classique employée dans la démographie et ce n'est que récemment qu'on a attiré l'attention sur l'insuffisance de cette approximation, (cf. Karmel [4]).Nous allons maintenant étudier la différence entre le nombre d'individus du sexe masculin et du sexe féminin.

Récemment Goodman [3] a étudié ce problème du point de vue de la théorie des processus stochastiques, et il donne une bibliographie assez complète dans son ouvrage. Nous avions étudié ce problème indépendamment en suivant, cependant, la même méthode et les résultats (53) donnés ici s'accordent avec les résultats obtenus par lui.

On considère une population comprenant des individus de deux types; disons type-I et type-II. Les individus de type-I donnent naissance à des individus de type-I et type-II tandis que les individus de type-II ne se reproduisent pas. Les individus de type-I et type-II correspondent respectivement aux individus du sexe féminin et du sexe masculin d'une population humaine.

Soit λp et λq $(p+q=1)$ les taux de natalité pour les individus de type-I et type-II respectivement et soit μ et μ' leurs taux de mortalité. Nous les définissons de même façon que dans la deuxième partie, tel que, par exemple, un individu de type-I en vie à l'instant t a la probabilité

$$\lambda_p(dt) + O(dt)$$

de donner naissance à un autre individu de type-I dans l'intervalle du temps $(t.\ t+dt)$. De plus, nous supposons que les taux de natalité et de mortalité ne dépendent ni de l'âge ni de l'époque t, et que la population se développe sous les mêmes conditions que pour le processus simple de naissance et de mort déjà considéré.

LES ÉQUATIONS FONDAMENTALES

Ces hypothèses faites, soit γ_t et δ_t les variables aléatoires qui désignent respectivement les nombres des individus de type-I et type-II dans la population à l'instant t et soit $P_{m,n}(t)$ la probabilité pour qu'à l'instant t la valeur de γ_t soit m et celle de δ_t soit n. Et supposons qu'à l'instant $t=0$ la variable aléatoire γ_t soit égale à 1 et δ_t soit égale à zéro. (L'extension au cas où les valeurs initiales de γ_t et δ_t sont d'autres que celles-ci est immédiate).

Dans ces conditions, la probabilité $P_{m,n}(t)$ satisfait aux équations différentielles

$$\frac{\partial P_{m,n}(t)}{\partial t} = -\left[m\lambda + m\mu + n\mu'\right] P_{m,n}(t) + \left[(m-1)\lambda p\right] P_{m-1,n}(t) + m\lambda q P_{m,n-1}(t) + (m+1)\mu P_{m+1,n}(t) + (n+1)\mu' P_{m,n+1}(t)$$

(49)
$$\frac{\partial P_{m,0}(t)}{\partial t} = -m(\lambda+\mu) P_{m,0}(t) + (m-1)\lambda p P_{m-1,0}(t) + (m+1)\mu P_{m+1,0}(t) + \mu' P_{m,1}(t)$$

$$\frac{\partial P_{0,n}(t)}{\partial t} = -n\mu' P_{0,n}(t) + \mu P_{1,n}(t) + \mu'(n+1) P_{0,n+1}(t)$$

$$\frac{\partial P_{0,0}(t)}{\partial t} = \mu P_{1,0}(t) + \mu' P_{0,1}(t)$$

avec les conditions initiales

(50)
$$P_{1,0}(0) = 1 \ ; \ P_{m,n}(0) = 0, \ m \neq 1 \ .$$

En prenant $P_{m,n}(t) = 0$ lorsque m ou n soit inférieur à zéro, la fonction génératrice des probabilités

$$\varphi(x,y,t) = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} P_{m,n}(t) x^m y^n$$

satisfait à l'équation aux dérivées partielles

$$(51) \qquad \frac{\partial \varphi}{\partial t} = (\lambda p x^2 + \lambda q x y - \lambda x - \mu x + \mu) \frac{\partial \varphi}{\partial x} + \mu'(1-y) \frac{\partial \varphi}{\partial y} .$$

La solution de cette équation partielle donnerait la loi du couple (γ_t, δ_t). Nous n'avons pas pu résoudre cette équation; cependant, tous les moments de la distribution peuvent être obtenus de la façon suivante.

Soit $M_{r,s}$ le moment (par rapport à l'origine), de l'ordre (r,s) du couple des variables aléatoires (γ_t, δ_t). En substituant $x = e^u$, $y = e^v$ dans l'équation (51) nous voyons que la fonction génératrice des moments

$$\Psi(u,v,t) = \sum_{0}^{\infty} \sum_{0}^{\infty} M_{r,s} \frac{u^r v^s}{r!\ s!}$$

satisfait à l'équation

$$(52) \qquad \frac{\partial \Psi}{\partial t} = \left[\lambda p e^u + \lambda q e^v - \lambda - \mu + \mu e^{-u}\right] \frac{\partial \Psi}{\partial u} + \mu'\left[e^{-v} - 1\right] \frac{\partial \Psi}{\partial v}$$

D'où on obtient

$$\frac{dM_{1,0}}{dt} = (\lambda p - \mu) M_{1,0}$$

$$\frac{dM_{0,1}}{dt} = \lambda q M_{1,0} - \mu' M_{0,1}$$

$$(53) \qquad \frac{dM_{2,0}}{dt} = (\lambda p + \mu) M_{1,0} + 2(\lambda p - \mu) M_{2,0}$$

$$\frac{dM_{1,1}}{dt} = (\lambda p - \mu - \mu') M_{1,1} + \lambda q M_{2,0}$$

$$\frac{dM_{0,2}}{dt} = 2 \lambda q M_{1,1} + \lambda q M_{1,0} + \mu' M_{0,1} - 2\mu' M_{0,2}$$

et ainsi de suite.

La solution de ces équations donne les moments et on en obtient,

$$E[\gamma_t] = e^{(\lambda p - \mu) t}$$

$$E[\delta_t] = \frac{\lambda q}{\lambda p - \mu + \mu'} \left[e^{(\lambda p - \mu) t} - e^{-\mu' t}\right]$$

$$\mathrm{Var}[\gamma_t] = \frac{\lambda p + \mu}{\lambda p - \mu} \quad e^{(\lambda p - \mu) t} \left[e^{(\lambda p - \mu) t} - 1\right]$$

$$\operatorname{Var}\left[\delta_t\right] = \frac{(\lambda p+\mu)\lambda^2 q^2}{(\lambda p-\mu+\mu')^2}\left[\frac{e^{2(\lambda p-\mu)t}}{(\lambda p-\mu)} + 2\frac{e^{(\lambda p-\mu-\mu')t}}{\mu'}\right.$$

$$(54)\qquad \left. - \frac{e^{-2\mu' t}}{\lambda p-\mu+2\mu'}\right] - \frac{2\lambda^2 q^2(\lambda p+\mu)e^{(\lambda p-\mu)t}}{\mu'(\lambda p-\mu)(\lambda p-\mu+2\mu')} + E\left[\delta_t\right]$$

$$\operatorname{Cov}\left[\gamma_t,\delta_t\right] = \frac{\lambda q(\lambda p+\mu)}{\mu'(\lambda p-\mu)(\lambda p-\mu+\mu')}\left[(\lambda p-\mu)e^{-\mu' t} + \mu' e^{(\lambda p-\mu)t} - (\lambda p-\mu+\mu')\right]e^{(\lambda p-\mu)t}$$

De (54) nous avons

$$(55)\qquad \frac{E[\gamma_t]}{E[\delta_t]} = \frac{(\lambda p-\mu+\mu')e^{(\lambda p-\mu)t}}{\lambda q\left[e^{(\lambda p-\mu)t}-e^{-\mu' t}\right]}$$

Si $\lambda p-\mu+\mu'$ est positif, ce rapport tend vers une constante $\left(=\frac{\lambda p-\mu+\mu'}{\lambda q}\right)$ lorsque t tend vers l'infini. Si nous supposons que les taux de mortalité μ et μ' sont égaux pour les individus de deux types

$$(56)\qquad \frac{E[\gamma_t]}{E[\delta_t]} = \frac{pe^{\lambda pt}}{q(e^{\lambda pt}-1)}$$

et nous voyons que dans ce cas ce rapport ne dépend pas de μ. De plus, dans ce cas, on a

$$(57)\qquad \lim_{t\to\infty}\frac{E[\gamma_t]}{E[\delta_t]} = \frac{p}{q}$$

C'est-à-dire que le rapport est, à la limite, égal au rapport des probabilités de naissances d'individus de type-I et de type-II.

LA LOI MARGINALE DE γ_t

Soit $P_m(t)$ la probabilité marginale pour que la valeur de γ_t à l'instant t soit égale à m. On a donc

$$P_m(t) = \sum_{n=0}^{\infty} P_{m,n}(t)\ .$$

En utilisant les relations (49) on trouve

$$(58)\qquad \frac{\partial P_m(t)}{\partial t} = -m(\lambda p+\mu)P_m(t) + (m-1)\lambda pP_{m-1}(t) + (m+1)\mu P_{m+1}(t) \qquad m > 1$$

$$\frac{\partial P_0(t)}{\partial t} = \mu P_1(t)$$

avec les conditions initiales

$$(59)\qquad P_1(0) = 1\ ;\quad P_m(0) = 0\ ,\quad m \neq 1\ .$$

En comparant (58) et (59) à (22) et (23) respectivement on constate que la variable aléatoire γ_t a pour la loi marginale la même loi qu'on a déjà considérée pour un processus simple de naissance et de mort. Ainsi les résultats obtenus dans la deuxième partie sont encore valables, et la probabilité $P_m(t)$ est donnée par (32) où on a maintenant

$$\frac{\xi}{\mu} = \frac{\mu}{\lambda p} = \frac{e^{(\lambda p-\mu)t}-1}{\lambda p e^{(\lambda p-\mu)t}-\mu}$$

PROBABILITÉ D'EXTINCTION D'UNE LIGNE DE DESCENDANCE

La probabilité d'extinction ne dépend que des individus de type-I. Or, le nombre d'individus de type-I croît (ou décroît) selon les résultats déjà obtenus pour un processus simple de naissance et de mort. Donc la probabilité d'extinction est

(60)
$$= 1 \quad \text{si} \quad \lambda p \leqslant \mu$$
$$= \frac{\mu}{\lambda p} \quad \text{si} \quad \lambda p > \mu$$

LA LOI DE PROBABILITÉ DU NOMBRE DES DESCENDANTS

La probabilité pour qu'un individu de type-I ait la durée de vie T, et ait eu pendant cette durée de vie (r+s) descendants de n'importe quel type est, après (45)

(61) $$\frac{1}{(r+s)!} (\lambda T)^{r+s} \quad \mu\, e^{-(\lambda+\mu)T}\, dT .$$

Et la probabilité pour que de ces (r+s) descendants r soient de type-I et s de type-II est, d'après la loi binome

$$C_{r+s}^{r} \quad p^r \; q^s .$$

Donc la probabilité $p_{r,s}$ pour qu'un individu de type-I ait eu r descendants de type-I et s de type-II (ni moins ni plus) avant sa mort est donnée par

$$P_{r,s} = \int_0^\infty C_{r+s}^{r} \frac{\mu}{(r+s)!} p^r q^s \quad (\lambda T)^{r+s} e^{-(\lambda+\mu)T} dT$$

c'est-à-dire

(62) $$P_{r,s} = C_{r+s}^{r} \; p^r \; q^s \; \left(1-\frac{\lambda}{\lambda+\mu}\right)\left(\frac{\lambda}{\lambda+\mu}\right)^{r+s} \qquad \begin{array}{l} r = 0,1,2,.. \\ s = 0,1,2,.. \end{array}$$

Encore la probabilité pour qu'un individu ait pendant la durée de sa vie r descendants de type-I est

$$p_r^{(1)} = \sum_{s=0}^{\infty} P_{r,s} = \left[1 - \frac{\lambda p}{\lambda p+\mu}\right] \left[\frac{\lambda p}{\lambda p+\mu}\right]^r \qquad r = 0,1,2,..$$

De même la probabilité pour qu'il ait s descendants de type-II est

$$p_s^{(2)} = \left[1 - \frac{\lambda q}{\lambda q + \mu}\right] \left[\frac{\lambda q}{\lambda q + \mu}\right]^s \qquad s = 0,1,2,\ldots$$

ESTIMATION DES PARAMÈTRES

Considérons un échantillon de n individus de type-I ayant la durée de vie

$$T_1, T_2, \ldots\ldots\ldots\ldots, T_n,$$

et supposons que chacun de ces n individus ait eu pendant la durée de sa vie

$$r_1, r_2, \ldots\ldots\ldots\ldots, r_n,$$

descendants de type-I, et

$$s_1, s_2, \ldots\ldots\ldots\ldots, s_n$$

descendants de type-II respectivement. Et soit

$$T = T_1 + T_2 + \ldots\ldots + T_n$$
$$R = r_1 + r_2 + \ldots\ldots + r_n$$
$$S = s_1 + s_2 + \ldots\ldots + s_n.$$

Alors, pour l'estimation $\hat{\lambda}$ et $\hat{\mu}$ de λ et μ respectivement on ne considère que le nombre total de descendants de chaque individu, et d'après les résultats de la deuxième partie, on a

(63) $$\hat{\lambda} = \frac{R + S}{T}, \qquad \hat{\mu} = \frac{n}{T}.$$

Et le couple $(\hat{\lambda}, \hat{\mu})$ suit asymptotiquement une loi de Laplace-Gauss avec les valeurs moyennes λ et μ respectivement et la matrice des covariances (48)

Pour l'estimation de p, c'est un problème d'estimation du paramètre d'une loi binome. Donc, l'estimation $\hat{p}$ de p est donnée par

(64) $$\hat{p} = \frac{R}{R + S}.$$

Il est évident que

(65) $$E\left[\hat{p}\right] = p$$
$$\mathrm{Var}\left[\hat{p}\right] = pq/R+S.$$

QUATRIÈME PARTIE

LA THÉORIE STOCHASTIQUE

Le processus généralisé de naissance et de mort

Jusqu'à présent nous avons supposé que les fonctions $\lambda(\alpha, t)$ et $\mu(\alpha, t)$ sont constantes et ne dépendent ni de l'âge α ni de l'époque t. Cette supposition, qui peut être suffisante pour le problème du développement des colonies de bactéries, est, cependant, un peu trop restrictives pour les problèmes de la démographie humaine. Par exemple, nous avons vu dans la deuxième partie que la probabilité de survie jusqu'à l'âge α, (voir formule (42) a la simple forme exponentielle $e^{-\mu\alpha}$. Or, la comparaison avec une table de survie montre l'insuffisance de cette forme de la fonction de survie. De même, il est évident que la probabilité de naissance λ dépend essentiellement de l'âge α de la mère. On peut supposer cette probabilité indépendante de t, mais on ne peut pas ignorer que la fécondité dépend de l'âge de la mère.

Le cas où λ et μ sont indépendants de l'âge mais dépendent de l'époque t a été traité en détail par Kendall [5] et Consael [1] . Ici nous laisserons de côté ce problème et examinerons seulement le cas où λ et μ dépendent de l'âge α mais non de l'époque t. La solution complète de ce problème est loin d'être achevée et nous nous contenterons de donner quelques résultats particuliers en utilisant la théorie des processus de Poisson généralisés. (cf. Fortet et Blanc-Lapierre [2]).

LA PROBABILITÉ DE SURVIE

On suppose maintenant qu'un individu encore en vie à un instant t quelconque et ayant alors l'âge α a la probabilité

$$\mu(\alpha)d\alpha + O(d\alpha)$$

de mourir dans l'intervalle $(t, t+d\alpha)$. Alors la probabilité de survie $p(\alpha)$, c'est-à-dire, la probabilité pour qu'un individu ait l'âge α, satisfait à l'équation

$$p(\alpha+d\alpha) = p(\alpha)\left[1 - \mu(\alpha)d\alpha - O(d\alpha)\right]$$

c'est-à-dire,

$$\frac{\partial p(\alpha)}{\partial \alpha} = -\mu(\alpha)p(\alpha) \quad ,$$

avec la condition initiale

$$p(0) = 1 \quad .$$

La solution de cette équation qui satisfait à la condition initiale est donnée par

(66) $$p(\alpha) = e^{-\int_0^\alpha \mu(a)\,da}.$$

Et la probabilité pour qu'un individu ait la durée de vie T, est

(67) $$\mu(T)e^{-\int_0^T \mu(a)\,da}\,dT \quad .$$

LES FONCTIONS DE FÉCONDITÉ

La probabilité pour qu'un individu en vie à l'instant t et ayant alors l'âge α donne naissance à un individu (de l'âge 0) dans l'intervalle $(t, t+d\alpha)$ est maintenant

$$\lambda(\alpha) + 0\,(d\alpha).$$

$\lambda(\alpha)$ correspond à la fonction de fécondité f(a) considérée dans la première partie.

La probabilité pour qu'un individu ne donne naissance à aucun individu entre les âges α_1 et α_2 est

(68) $$e^{-\int_{\alpha_1}^{\alpha_2} \lambda(a)\,da}$$

Donc, si $p(\alpha_1, \alpha_2, \ldots\ldots, \alpha_n)\,d\alpha_1\,d\alpha_2 \ldots\ldots d\alpha_n$ représente la probabilité d'avoir un enfant dans chacun des n intervalles d'âge

$$(\alpha_1, \alpha_1+d\alpha_1),\quad (\alpha_2, \alpha_2+d\alpha_2),\ \ldots\ldots,\ (\alpha_n, \alpha_n+d\alpha_n)$$

où

$$\alpha_1 < \alpha_2 < \ldots\ldots\ldots\ldots, < \alpha_n,$$

on a

(69) $$\begin{aligned} & p(\alpha_1, \alpha_2, \ldots\ldots, \alpha_n)\,d\alpha_1\,d\alpha_2 \ldots\ldots d\alpha_n \\ & = \left[e^{-\int_0^{\alpha_1} \lambda(a)\,da}\,\lambda(\alpha_1)\,d\alpha_1\right]\left[e^{-\int_{\alpha_1}^{\alpha_2} \lambda(a)\,da}\,\lambda(\alpha_2)\,d\alpha_2\right] \ldots\ldots \\ & \qquad \left[e^{-\int_{\alpha_{n-1}}^{\alpha_n} \lambda(a)\,da}\,\lambda(\alpha_n)\,d\alpha_n\right] \\ & = \lambda(\alpha_1)\,\lambda(\alpha_2) \ldots\ldots \lambda(\alpha_n)\ e^{-\int_0^{\alpha_n} \lambda(a)\,da}\,d\alpha_1\,d\alpha_2 \ldots\ldots d\alpha_n \end{aligned}$$

Soit maintenant

$$P_n(\alpha)$$

la probabilité d'avoir n enfants entre les âges 0 et α . $P_n(\alpha)$

satisfait aux équations

$$\frac{\partial P_n(\alpha)}{\partial \alpha} = -\lambda(\alpha) P_n(\alpha) + \lambda(\alpha) P_{n-1}(\alpha) \qquad n \geqslant 1$$

$$\frac{\partial P_0(\alpha)}{\partial \alpha} = -\lambda(\alpha) P_0(\alpha)$$

avec les conditions initiales

$$P_0(0) = 1 \quad ; \quad P_n(0) = 0 \ , \quad n > 0 \ .$$

Donc, on a

(70) $$P_n(\alpha) = \frac{1}{n!}\left[\int_0^{\alpha} \lambda(a)\,da\right]^n \times \ e^{-\int_0^{\alpha} \lambda(a)\,da}$$

Si maintenant $\pi_n(\alpha)d\alpha$ représente la probabilité d'avoir le nième enfant à l'âge α – ($\pi_n(\alpha)$ correspond à la fonction $f_n(a)$ de la première partie) – on a

(71) $$\pi_n(\alpha) = P_{n-1}(\alpha)\ \lambda(\alpha)$$

$$= \frac{1}{(n-1)!}\left[\int_0^{\alpha} \lambda(a)\,da\right]^{n-1} e^{-\int_0^{\alpha} \lambda(a)\,da} \quad \lambda(\alpha)$$

LE NOMBRE DE DESCENDANTS

Soit p_r la probabilité qu'un individu ait eu r enfants pendant la durée de sa vie. La probabilité pour qu'un individu ait la durée de vie T est

$$\mu(T) e^{-\int_0^T \mu(a)\,da} \quad dT \ .$$

Et la probabilité d'avoir r enfants entre les âges 0 et T est

$$\frac{1}{r!}\left[\int_0^T \lambda(a)\,da\right]^r e^{-\int_0^T \lambda(a)\,da} \ .$$

Donc

$$p_r = \int_0^{\infty} \frac{1}{r!}\mu(T)\left[\int_0^T \lambda(a)\,da\right]^r e^{-\int_0^T \left[\lambda(a)+\mu(a)\right] da} \ dT$$

$$= \int_{T=0}^{\infty} \frac{1}{r!}\left\{L(T)\right\}^r e^{-\left[L(T)+M(T)\right]} \ d\left[M(T)\right]$$

où

$$\int_0^{\alpha} \lambda(a)\,da = L(\alpha) \quad , \quad \int_0^{\alpha} \mu(a)\,da = M(\alpha) \ .$$

Si L(T)/M(T) est constant (égal à k), on a

$$p = \int_{T=0}^{\infty} \frac{1}{r!} k^r \left[M(T)\right]^r \; e^{-M(T)\left[1+k\right]} \, dM(T)$$

(72)

$$= \left(1 - \frac{k}{k+1}\right)\left(\frac{k}{k+1}\right)^{r} \text{(1)} \qquad r = 0,1,2,\ldots..$$

Ainsi, même dans le cas où λ et μ dépendent de l'âge, la probabilité p_r a encore la même forme géométrique de (46) si le rapport λ/μ est constant.

RÉPARTITION PAR AGE

Nous désignons encore par $d\gamma(\alpha,t)$ le nombre d'individus dont l'âge à l'instant t est compris entre les âges α et $\alpha+d\alpha$, et nous faisons les mêmes hypothèses sur $d\gamma(\alpha,t)$, à savoir que $d\gamma(\alpha,t)$ suit asymptotiquement la loi de Poisson dont l'espérance mathématique et la variance sont approximativement égales à $\delta(\alpha,t)d\alpha$.

Si $\alpha < t$, le nombre d'individus dans le groupe d'âges $(\alpha, \alpha+d\alpha)$ dépend seulement du nombre d'individus nés à l'époque $t-\alpha$ et en vie à l'instant t. Nous supposons qu'à l'instant t=0 la population comprenait un seul individu (l'ancêtre) de l'âge 0.

La probabilité que l'ancêtre soit en vie à l'instant $t-\alpha$ est

$$e^{-\int_0^{t-\alpha} \mu(a)da} ,$$

et la probabilité qu'il donne naissance à un individu à l'instant $t-\alpha$ est

$$\lambda(t-\alpha)d\alpha.$$

Donc l'espérance mathématique du nombre de naissances à l'instant $t-\alpha$ et dues à l'ancêtre est

$$\lambda(t-\alpha)e^{-\int_0^{t-\alpha} \mu(a)\,da}\, d\alpha$$

L'espérance mathématique du nombre d'individus dont l'âge à l'instant $t-\alpha$ est compris entre les âges $(\beta, \beta+d\beta)$ est

$$\delta(\beta, t-\alpha)d\beta , \quad \beta < t-\alpha .$$

---

(1) Pour être valable le résultat impose les conditions suivantes :

(i) $m(\alpha) = 0$ pour $\alpha = 0$
(ii) $m(\alpha) = \infty$ pour $\alpha = \infty$
(iii) $m(\alpha) > 0$ pour $\alpha > 0$

Donc l'espérance mathématique du nombre de naissances à l'instant $t-\alpha$ et dues aux individus autres que l'ancêtre est

$$\left[\int_0^{t-\alpha} \lambda(\beta)\, \delta\, (\beta, t-\alpha)\, d\beta\right] d\alpha\ .$$

L'espérance mathématique du nombre total des naissances à l'instant $t-\alpha$ est

$$\left[\lambda(t-\alpha)\, e^{-\int_0^{t-\alpha} \mu(a)\, da} + \int_0^{t-\alpha} \lambda(\beta)\delta(\beta, t-\alpha)\, d\beta\right] d\alpha\ .$$

La probabilité pour qu'un individu né à l'instant $t-\alpha$ soit en vie à l'instant t est

$$e^{-\int_0^{\alpha} \mu(a)\, da}$$

Donc l'espérance mathématique du nombre d'individus en vie parmi ceux nés à l'instant $t-\alpha$ est

$$e^{-\int_0^{\alpha} \mu(a)\, da}\left[\lambda(t-\alpha)\, e^{-\int_0^{t-\alpha} \mu(a)\, da} + \int_0^{t-\alpha} \lambda(\beta)\delta(\beta, t-\alpha)\, d\beta\right] d\alpha$$

$$= \delta(\alpha, t)\, d\alpha\ ,$$

c'est-à-dire,

$$(73) \qquad \delta(\alpha, t) = e^{-\int_0^{\alpha} \mu(a)\, da}\left[\lambda(t-\alpha)\, e^{-\int_0^{t-\alpha} \mu(a)\, da} + \int_0^{t-\alpha} \lambda(\beta)\delta(\beta\ , t-\alpha)\, d\beta\right].$$

Ainsi nous avons pour $\delta(\alpha, t)$ une équation intégrale. Si l'on suppose que seul α dépend de l'âge et que μ est constant, cette équation se réduit à la forme trouvée par Kendall [6] .

CONCLUSION

Dans cette étude nous avons essayer d'utiliser la théorie des processus stochastiques en démographie humaine. Notre but était de montrer comment l'étude des problèmes démographiques peut être faite d'une façon plus satisfaisante par la méthode probabiliste et de montrer la correspondance étroite qui existe entre les résultats de la théorie déterministe et ceux de la théorie stochastique. Ainsi, nous n'avons pas donné les démonstrations rigoureuses et complètes pour tous les résultats. Nous nous sommes contentés parfois d'indiquer seulement les résultats intéressants en conservant en même temps un raisonnement aussi simple que possible.

Signalons ici certaines des conclusions qu'on peut tirer de cette étude élémentaire.

La théorie stochastique n'ayant, à la base que deux fonctions λ et μ, présente une simplification assez importante du problème en comparaison de la méthode classique. Nous avons vu qu'à partir de ces deux fonctions on peut obtenir presque tous

les résultats de la démographie mathématique. De plus, nous avons vu que pour le problème du nombre de descendants et le problème connexe d'extinction d'une population aucune autre hypothèse n'est nécessaire. La méthode d'estimation de λ et μ dans le cas où λ et μ sont constants peut être généralisée pour le cas où λ et μ dépendent de l'âge et peut servir pour une nouvelle définition des taux de natalité et de mortalité.

On peut aussi généraliser la méthode adoptée dans la troisième partie pour tenir compte de la nuptialité et d'autres facteurs afin d'obtenir une représentation plus proche de la réalité.

BIBLIOGRAPHIE

1 R. Consael, "Sur une généralisation du processus de Polya", Acad.Roy.Belgique Bull.Cl.Sci(5),34(1948),863-76.

2 R. Fortet et A. Blanc-Lapierre, Théorie des Fonctions Aléatoires, Paris, Masson et Cie, édit. (1953)

3 Leo A. Goodman, "Population growth of sexes," Biometrics, 9 (1953), 212-25.

4 P.H. Karmel, "The relations betwen male and female reproduction rates,"Population Studies, 1(1947), 249-74.

5 D.G. Kendall, "On the generalised birth and death process," Ann. Math. Stat., 19 (1948), 1-15.

6 " "Stochastic processes and population growth, Jour.Roy.Stat.Soc.,Ser B, 11(1949),230-64.

7 " "Random fluctuations in the age distribution of a population whose development is controlled by the simple birth and death process," Jour.Roy.Stat.Soc.,Ser B, 12(1950), 278-85.

8 " "Les processus stochastiques de croissance en biologie," Ann. Inst. H.Poincaré, t.13 (1952), facs.1, 43-108.

9 V.A. Kostitzin, Biologie Mathématique, Paris, A. Colin édit. (1939).

10 A.J. Lotka, Théorie Analytique des Associations Biologiques, Actual. Scient. n°780, Paris, Hermann édit.(1939)

11 P.A.P. Moran, "Estimation methods for evolutive processes," Jour.Roy.Stat.Soc., Ser B, 13(1951), 141-46.Voir aussi Jour.Roy.Stat.Soc.,Ser B,15(1953),241-45.

12 E.C. Rhodes,"Population mathématics," Jour.Roy. Stat. Soc. , 103(1940), 61-89, 218-45 et 362-87.

13 J.F. Steffensen, "Deux problèmes du calcul des probabilités" Ann.Inst. H. Poincaré, 3(1933), 319-44.

TABLE DES MATIÈRES

CONSIDÉRATIONS STATISTIQUES SUR LES POLYGONES ET LES POLYÈDRES

par

Matthias MATSCHINSKI,
Ancien Professeur d'Université

I. PRÉLIMINAIRES. STATISTIQUE DE POLYGONES

§ 1 – Dans l'histoire de la géométrie on a essayé, et on a accompli, plusieurs modes d'extension et de généralisation. Toutes ces tentatives sont bien connues et il est peu intéressant de les rappeler: géométries non-euclidienne, non-desarguienne, etc.. Cependant il reste toujours un nombre suffisant de chemins non encore essayés. Un de ceux-ci, et nous l'essayons ici, est l'extension du concept du polygone (polyèdre, etc..) introduisant les polygones au nombre non entier de côtés (les polyèdres au nombre non entier d'arêtes, etc...). Jusqu'ici, d'après notre connaissance, on n'a considéré que les figures au nombre entier d'angles, de sommets, de côtés, d'arêtes, de faces, etc.. Nommons tous ces nombres, les <u>nombres caractéristiques</u> de la figure. L'introduction des nombres caractéristiques non-entiers peut être réalisé à l'aide d'un système d'axiomes. Mais nous préférons, pour premier essai d'une telle géométrie, choisir les considérations statistiques qui nous amènent aux nombres caractéristiques fractionnaires ou même irrationnels. De plus, une géométrie construite sur la base de la statistique sera tout de suite directement applicable aux nombreux problèmes pratiques, conservant en même temps toute signification de nouvelle géométrie abstraite.

§ 2 – Ainsi, en considérant les nombres caractéristiques moyens correspondant à une série de figures, nous établissons l'extension cherchée dans sa forme primitive. Mais si les séries de figures ne sont pas caractérisées par quelque trait spécifique géométrique, les moyennes sont complètement arbitraires, et ne dépendent que de la distribution des figures dans la série choisie. Ce ne sera que le sujet d'études dans le sens de la statistique générale sur les probabilités, les distributions moyennes etc... d'une série fortuite. Pour introduire une liaison géométrique entre les figures de la série, supposons que ces figures couvrent sans lacunes un espace quelconque, plan ou non-plan, fermé ou non-fermé, etc.. (– <u>condition de base</u>). Les séries de polyèdres et de polygones ont été déjà traitées par plu-

sieurs auteurs; par exemple H. Minkowski[1] , ou, pour citer le plus ancien: J. Steiner[2] . Mais toutes les publications précédentes n'introduisent pas les considérations statistiques.Toutes elles appartiennent à la géométrie classique, non-statistique,et elles ne sont citées ici que du point de vue historique. Dans les séries "statistiques" déterminées par la condition de base mentionnée, les nombres caractéristiques ne sont pas complètement arbitraires, mais sont liés entre eux, ne laissant que quelques uns entièrement arbitraires. Par exemple, les polygones dans un plan euclidien sont caractérisés en moyenne, par un nombre moyen de côtés d'une figure ($\bar{n}$) et un nombre moyen de côtés s'entrecroisant en un point($\bar{\mu}$). On peut construire une série ayant n'importe quelle valeur numérique, entière ou non-entière, de nombre moyen de côté ($\bar{n}$) et une autre série ayant n'importe quelle valeur de μ (sur la limite de la variation de n et de μ voir III § 1). Cependant, dans le premier cas ($\bar{n}$ arbitraire fixé) $\bar{\mu}$ ne peut prendre qu'une seule valeur numérique et vice-versa. Ainsi, aux exemples classiques des hexagones, des carrés et des triangles, couvrant le plan sans lacunes, on peut ajouter aussi les pentagones couvrant le plan sans lacunes sous la condition que $\bar{\mu}$ est égal à 3 1/3, c'est-à-dire sous la condition qu'en chaque point s'entrecroisent 3 1/3 lignes. Cela peut être réalisé de plusieurs façons, par exemple de façon qu'en un tiers des points d'entrecroisement se touchent quatre côtés, tandis que dans les deux autres tiers ne se touchent que trois lignes . Cette liaison prend dans le cas du plan euclidien la forme:

$$\frac{1}{\bar{n}} + \frac{1}{\bar{\mu}} = \frac{1}{2} \qquad (1)$$

comme ce sera démontré par la suite.

§ 3 - Ce fait peut être trouvé soit comme un fait de la géométrie expérimentale, soit en partant de considérations théoriques. Nous reviendrons à la méthode expérimentale dans le III§4; ici, nous considérons les raisons théoriques. Dans notre note[3] nous avons indiqué la possibilité de démontrer la relation (1) à l'aide de considérations probabilistes. Pour ne pas répéter cette démonstration, nous en indiquons ici une autre si simple qu'elle peut apparaître presque banale.

Supposons que notre série de polygones forme un polyèdre fini ou infini ayant e sommets, k arêtes et f faces. Il est essentiel que les grandeurs e, k, f soient, pour le cas de la géométrie statistique d'un espace à deux dimensions, non statistique, contrairement aux grandeurs, $\bar{n}, \bar{\mu}$ substantiellement statistiques. Pour le polyèdre qui est formé par la série donnée, la formule d'Euler:

$$e + f = k + 2 \qquad (2)$$

est valable. Comptons maintenant le nombre de côtés (n) pour chaque polygone entrant dans la série et formons la somme de tous ces côtés: $\sum n$. Chaque côté des polygones touchant deux sommets, la somme obtenue n'est que le nombre doublé des arêtes (2 k):

$$2k = \sum n = f \frac{\sum n}{f} = f \cdot \bar{n} \qquad (3)$$

d'après le concept de la moyenne. En toute analogie on trouve

$$2k = \Sigma\mu - e\frac{\Sigma\mu}{e} - e\cdot\bar{\mu} \quad (4)$$

Substituant (3) et (4) dans (2) on obtient:

$$\frac{1}{\bar{n}} + \frac{1}{\bar{\mu}} = \frac{1}{2} + \frac{1}{k}, \quad (5)$$

formule générale pour n'importe quelle série de figures. Si nous considérons une série dans un plan euclidien, le polyèdre auxiliaire, introduit dans la démonstration de la formule (5), sera infiniment grand, le nombre k aussi, et la formule (5) se réduira à (1).

Pour simplifier l'écriture, introduisons le symbole R a pour exprimer $\frac{1}{a}$. Avec ce symbole la formule (1) s'écrira:

$$R\,\bar{n} + R\,\bar{\mu} = R\,2 + R\,k \quad (6)$$

§ 4 – La formule (1) [ou (6)] n'est pas aussi élémentaire qu'on peut le supposer à première vue. Il existe une autre formule presque évidente qui rappelle (6), mais dont le sens est

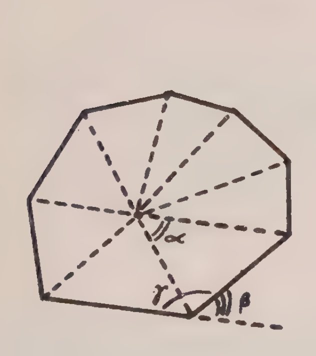
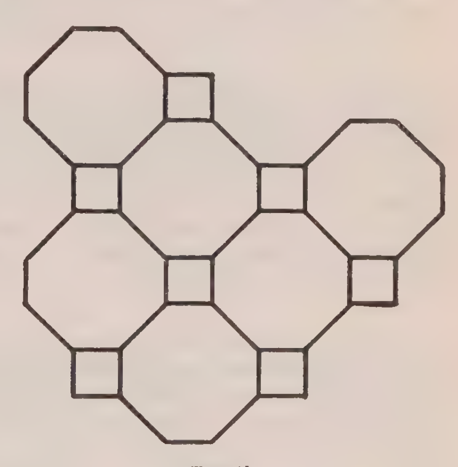


Fig. 1a


Fig. 1b

essentiellement différent de (6). Considérons la figure 1a. Il est évident que

$$\gamma + \beta = \pi \quad (7)$$

pour chaque triangle, élément de polygone. La formule (7) étant linéaire, on trouve de même

$$\bar{\gamma} + \bar{\alpha} = \pi \quad \left(\text{parce que } \bar{\alpha} = \bar{\beta}\right) \quad (8)$$

pour tous les éléments d'un polygone et, enfin, pour tous les polygones. D'après le sens de grandeur n – nombre de côtés d'un

seul polygone et μ - nombre de côtés se touchant en un sommet,on a en moyenne:

$$\overline{\left(\frac{2\pi}{\mu}\right)} + \overline{\left(\frac{2\pi}{n}\right)} = \pi \qquad (9)$$

d'où découle

$$\overline{R\mu} + \overline{Rn} = R2 \, , \qquad (10)$$

qui est la dite formule élémentaire. Mais on ne doit pas oublier que $R\bar{n} \neq \overline{Rn}$,etc. On voit que les formules (10) et (6) ont des formes très proches à première vue. Cependant leur sens est très différent; la formule (6) contient en effet les valeurs moyennes qui nous sont nécessaires pour la construction de la géométrie statistique de polygones, tandis que la formule (10) ne contient que les moyennes des valeurs réciproques qui n'ont aucune liaison avec $\bar{n}$, $\bar{\mu}$ comme il sera démontré plus bas. Voici un exemple illustrant la différence entre les formules (6) et (10) - voir figure 1b.

$$\begin{gathered}\bar{n} = \frac{4+8}{2} = \bar{n} = 6 \, ; \; \bar{\mu} = \mu = 3 \, ; \; \frac{1}{\bar{n}} + \frac{1}{\bar{\mu}} = \frac{1}{6} + \frac{1}{3} = \frac{1}{2} \, ; \\ \overline{\left(\frac{1}{n}\right)} = \frac{1}{2}\left(\frac{1}{8}+\frac{1}{4}\right) = \frac{3}{16} \quad ; \quad \overline{\left(\frac{1}{\mu}\right)} = \frac{1}{2}\left(\frac{1}{4}+\frac{3}{8}\right) = \frac{5}{16} \quad ; \quad \overline{\left(\frac{1}{n}\right)} + \overline{\left(\frac{1}{\mu}\right)} = \frac{3}{16} + \frac{5}{16} = \frac{1}{2} \, .\end{gathered} \qquad (11)$$

Que cet exemple soit constitué de polygones réguliers n'est pas important pour la théorie développée. Au contraire, celle-ci suppose les séries des polygones complètement arbitraires, réguliers et non-réguliers, disposés en ordre ou disposés en désordre. Si nous avons choisi ici, et choisirons plusieurs fois ci-après, les séries des polygones réguliers et disposés régulièrement, c'est seulement pour la raison que l'estimation des probabilités et, par conséquent, des moyennes, est beaucoup plus facile dans de tels cas.

§ 5 - La formule générale (5) caractérisant la statistique des polygones sur une surface quelconque est représentée sur les fig. 2 et 3. Pour cette représentation graphique il est commode de récrire cette formule (5) en tenant compte de (3) sous la forme:

$$\bar{n} = \frac{2\bar{\mu}}{\bar{\mu}-2}\left(1 - \frac{2}{f}\right) \qquad (12)$$

Sur l'axe horizontal de la fig. 2 sont portées les valeurs réciproques de f - fractions de l'angle sphérique total qui correspondent à chaque polygone; l'axe vertical représente les nombres de côtés de polygones. Comme on le voit d'après la formule (12) ces deux grandeurs ($\frac{1}{f}$ et $\bar{n}$) sont liées linéairement, si $\bar{\mu}$ est constante. Ces lignes droites correspondant aux différentes valeurs de $\bar{\mu}$, sont indiquées sur la fig. 2 par des droites pointillées; elles ont un point commun A. Leurs intersections avec l'axe de $\bar{n}$ donnent les nombres de côtés de polygones correspondant au cas d'un plan euclidien: nous voyons ici les valeurs connues n = 3 pour la ligne correspondant à μ = 6, n = 4

pour la ligne correspondant à $\mu = 4$, etc..., ainsi que la valeur $n = 3\ 1/3$ pour la ligne correspondant à $\mu = 5$ et $n = 5$ pour la ligne correspondant à $\mu = 3\ 1/3$.

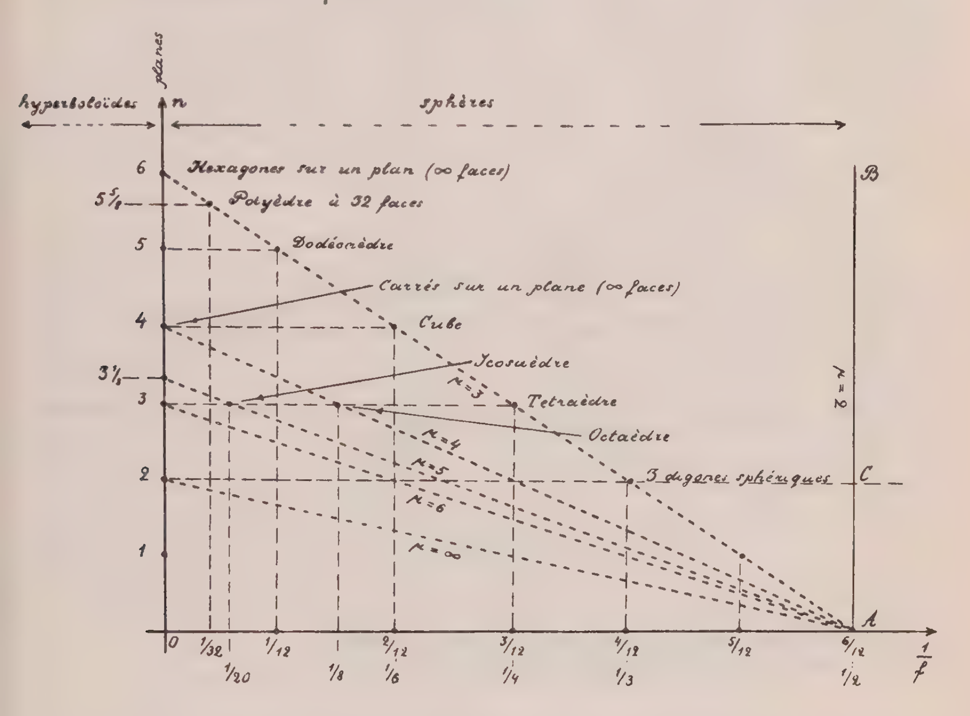


Fig. 2

Pratiquement la ligne limite est la ligne $\mu = 6$, mais du point de vue formel, on peut même considérer la ligne $\mu = \infty$. La ligne $\mu = 2$ est une ligne verticale AB; elle correspond aux polygones couplés, formellement représentables par les équateurs d'une sphère ou par un plan ayant deux côtés et limités par un polygone, un dièdre. La ligne droite 2C représente, soit les prismes infinis, soit les méridiens d'une sphère. L'emplacement de quelques formes classiques est indiqué sur la figure. Cependant, on ne doit pas oublier que chaque point, même l'emplacement d'une forme classique, correspond en même temps à une infinité de séries de polygones irréguliers ayant mêmes $\bar{n}$ et $\bar{\mu}$ que n et μ de cette figure classique. Comme les formules du § 3 le montrent, à ces deux valeurs données, celles de e, k, f correspondent univoquement. La dépendance entre $\bar{n}$ et $\bar{\mu}$ pour f donné n'est pas linéaire comme celle entre $\bar{n}$ et $\frac{1}{f}$; on voit que $\bar{\mu} = \varphi\,(\bar{n})$ d'après la fig. 3. Tout cela est très élémentaire, quoique n'ayant jamais été dit dans le sens statistique ; les difficultés commencent dans le cas de trois et de plusieurs dimensions.

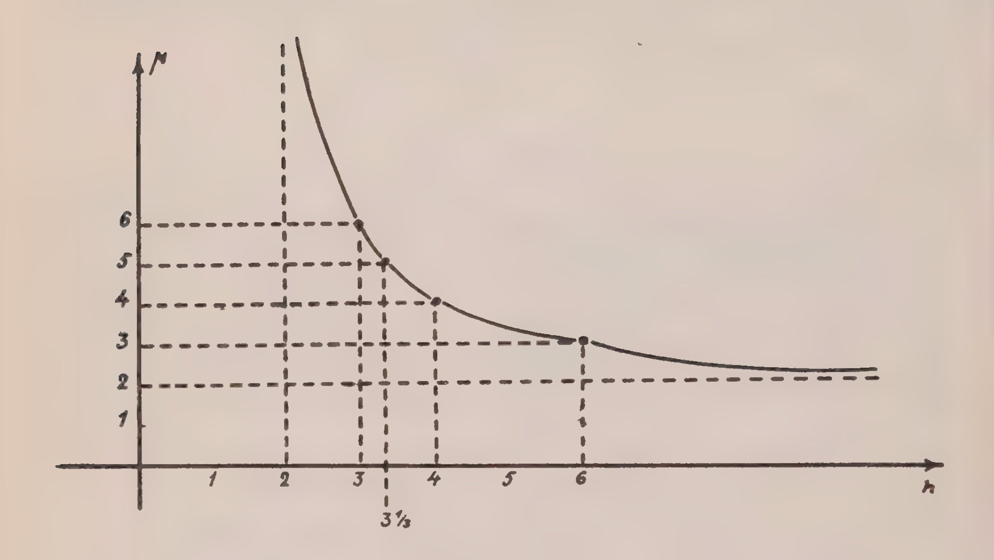


Fig. 3

II. STATISTIQUE DE POLYÈDRES ET DE FIGURES DE QUATRE DIMENSIONS OU PLUS

§ 1 - La difficulté essentielle de la statistique des polyèdres et des figures à plusieurs dimensions (*) est de même espèce que la difficulté bien connue de la statisation en hydrodynamique: l'impossibilité de calculer la valeur moyenne d'un produit à partir des valeurs moyennes de ses sous-multiples. Nous appellerons brièvement cette difficulté: "difficulté hydrodynamique". Par exemple, pour les polyèdres classiques, on a

$$e+f=k+2 \; ; \quad e\mu = 2k \; ; \quad fn=2k. \tag{13}$$

En considérant une série de polyèdres, on doit passer aux moyennes; cependant, si la première formule de (13) se transforme, sans difficultés, en $\bar{e} + \bar{f} = \bar{k} + 2$, il n'en est pas de même des deuxième et troisième formules de (13). Comme on le démontrera ci-après, le procédé qui permet d'éviter cette difficulté est très simple, à savoir: pour la statistique de figures à N dimensions on doit les considérer comme les éléments d'une superfigure à N+1 dimensions, considérer les grandeurs caractéristiques de cette superfigure comme les grandeurs non-statistiques et, enfin, revenir à la statistique de figures à N dimensions.

*) - polytopes, d'après la terminologie répandue, voir p.e. T.N. Schoute[4]

§ 2 - Expliquons-nous sur l'exemple de séries de polyèdres. Pour abréger nous désignons la superfigure sus-mentionnée comme le "monde". Dans le cas de polygones traités statistiquement, ce "monde" a été une surface quelconque fermée ou non-fermée, plane ou non plane, formant en général un polyèdre non-statistique. Dans le cas de polyèdres traités statistiquement le "monde" sera un espace quelconque, fermé ou non-fermé, ayant une courbure ou ne l'ayant pas, formant un superpolyèdre. La généralisation de la formule d'Euler. pour le cas de quatre dimensions, comme on le sait, est (voir aussi II § 5)

$$E + F = K + M \qquad (14)$$

où E, K, F et M sont les nombres respectifs de tous les sommets, de toutes les arêtes, de toutes les faces, de tous les polyèdres. Conservons aux lettres μ, n, k, f, e le même sens que dans la partie I. Introduisons encore: $\overline{\mu}_1$ - nombre moyen de toutes les lignes arêtes se touchant en un point; $\overline{\mu}_2$ - nombres moyens de toutes les surfaces (faces) se touchant en un point; $\overline{\mu}_3$ - nombres moyens des polygones (corps) se touchant en un point, enfin λ nombres de surfaces (faces) se touchant par une arête, ou ce qui revient au même, nombres des polygones (corps) se touchant par une arête. Au lieu d'employer, comme nous venons de le faire, l'expression "nombre moyen des éléments se touchant en un point" on peut introduire l'expression "le nombre des éléments se trouvant statistiquement autour d'un sommet (point)".

D'après le sens de valeurs moyennes et en analogie complète avec la formule (3), on a

$$2K = \Sigma\mu_1 = E \cdot \frac{\Sigma\mu_1}{E} = \overline{\mu}_1 \cdot E \;\; ; \;\; 2F = \Sigma f = M\frac{\Sigma f}{M} = \overline{f} \cdot M$$

$$\Sigma\mu_3 = E\frac{\Sigma\mu_3}{E} = \overline{\mu}_3 E = \Sigma e = M\frac{\Sigma e}{M} = \overline{e}M, \text{ analogiquement :} \qquad (15)$$

$$\overline{n}F = \overline{\mu}_2 E = \overline{k}M = \overline{\lambda}K.$$

Sur ces formules on doit faire une remarque importante. En divisant pour former les moyennes, une fois par E, l'autre fois par K, etc.. on n'introduit aucun arbitraire; le choix du diviseur est déterminé par le caractère de la grandeur dont on calcule la moyenne. Par exemple, en faisant les moyennes de e,k,f, on divise par M - nombre de polyèdres, parce que ces grandeurs se rapportent à un polyèdre; de même, pour calculer les moyennes de μ_1, μ_2, μ_3, on divise par E - nombre de sommets (points)parce que les grandeurs se rapportent à un point, etc.

§3 - Il existe un dualisme entre les caractéristiques statistiques des figures pour lesquelles on construit la statistique et les caractéristiques statistiques d'un point. De même une analogie totale existe entre les formules pour les figures-éléments et les formules pour les points.Dans le cas de polygones on peut remplacer les polygones par les sommets et vice-versa d'après le schéma:

$$\overline{\mu} \rightleftarrows \overline{n} \quad ; \quad e \rightleftarrows f \quad ; \quad k \rightleftarrows k \quad ;$$

et à cause de la symétrie des formules parvenir aux mêmes expressions. Dans le cas de polygones il existe la même réciprocité qui peut être brièvement exprimée par le schéma

$$\begin{matrix} M, & K, & \bar{\mu}, & \bar{n}, & \bar{e}, & \bar{k}, & \bar{f} \\ \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow \\ E, & F, & \bar{\mu}, & \bar{\lambda}, & \bar{\mu}_3, & \bar{\mu}_2, & \bar{\mu}_1 \end{matrix} \qquad (16)$$

si simple qu'il ne nécessite pas d'explication. Ainsi, la statistique de polygones étant établie, les formules

$$\bar{e}+\bar{f}=\bar{k}+2 \quad ; \quad \bar{\mu}_1+\bar{\mu}_3=\bar{\mu}_1+2 \; ; \qquad (17)$$

doivent aussi être considérées comme établies en prenant les moyennes des formules linéaires (voir § 1 de la partie II). En combinant les formules (15) avec (17) on trouve le système:

$$\begin{gathered} 2\bar{k}=\bar{e}\bar{\mu}=\bar{f}\cdot\bar{n} \; ; \; R\bar{n}+R\bar{\mu}=R2+R\bar{k} \; ; \; 2\bar{\mu}_2=\bar{\mu}\bar{\mu}_3=\bar{\lambda}\bar{\mu}_1 \; ; \\ R\bar{\lambda}+R\bar{\mu}=R2+R\bar{\mu}_2 \; ; \; 2F=\bar{f}M \; ; \; 2K=\bar{\mu}_1 E \; ; \bar{\lambda}K=\bar{k}M , \end{gathered} \qquad (18)$$

qui, avec une relation trigonométrique[3] forment un système complet des équations de la statistique des polyèdres. De manière un peu étonnante on obtient que le signe de moyenne et la multiplication sont permutables

$$\overline{e\mu}=\bar{e}.\bar{\mu} \quad ; \quad \overline{fn}=\bar{f}.\bar{n} \quad ; \quad \overline{\mu\mu_3}=\bar{\mu}.\bar{\mu}_3 \quad ; \quad \overline{\lambda\mu_1}=\bar{\lambda}.\bar{\mu}_1 \qquad (19)$$

Mais on ne peut pas prévoir ce résultat d'avance et pour chaque statistique des figures à N dimensions, on doit de nouveau appliquer le procédé indiqué en ayant recours à l'espace à N+1 dimensions, comme nous l'avons fait pour le cas des polyèdres.

Les formules (18) peuvent être présentées sous des formes très différentes. On peut citer les suivantes: forme où les grandeurs non-statistiques E,K,F,M, sont supposées être connues

$$\begin{gathered} \bar{\lambda}=\frac{2\bar{\mu}}{\bar{\mu}-2}\left(1-\frac{E}{K}\right) ; \; \bar{\mu}_2=\frac{2\bar{\mu}}{\bar{\mu}-2}\left(\frac{K}{E}-1\right) ; \; \bar{\mu}_1=\frac{2K}{E} ; \; \bar{\mu}_3=\frac{2\bar{\mu}_2}{\bar{\mu}}=\frac{4}{\bar{\mu}-2}\left(\frac{K}{E}-1\right), \\ \bar{n}=\frac{2\bar{\mu}}{\bar{\mu}-2}\left(1-\frac{M}{F}\right) ; \; \bar{k}=\frac{2\bar{\mu}}{\bar{\mu}-2}\left(\frac{F}{M}-1\right) ; \; \bar{f}=\frac{2F}{M} ; \; \bar{e}=\frac{4}{\bar{\mu}-2}\left(\frac{F}{M}-1\right), \end{gathered} \qquad (20)$$

et la forme où E,K,F,M, sont supposés être éliminés

$$\begin{gathered} \bar{\lambda}=\frac{2\bar{\mu}}{\bar{\mu}_1}\left(\frac{\bar{\mu}_1-2}{\bar{\mu}-2}\right) ; \; \bar{\mu}_2=\bar{\mu}\left(\frac{\bar{\mu}_1-2}{\bar{\mu}-2}\right) ; \; \bar{\mu}_3=2\left(\frac{\bar{\mu}_1-2}{\bar{\mu}-2}\right) ; \\ \bar{n}=\frac{2\bar{\mu}}{\bar{f}}\left(\frac{\bar{f}-2}{\bar{\mu}-2}\right) ; \; \bar{k}=\bar{\mu}\left(\frac{\bar{f}-2}{\bar{\mu}-2}\right) ; \; \bar{e}=2\left(\frac{\bar{f}-2}{\bar{\mu}-2}\right) . \end{gathered} \qquad (21)$$

Nous reviendrons à ces formules dans la partie III. Il suffit ici de souligner que dans (20) il n'entre que les rapports de grandeurs non-statistiques, et il n'entre pas de valeurs absolues; tandis que dans le cas de polygones, les valeurs statistiques ont été déterminées par les valeurs non-statistiques [voir les formules (5) et (12)] .Cela tient à la différence essentielle bien connue entre les espaces à nombres pairs et impairs de dimensions.

§ 4- On voit alors que pour la construction de la statistique de polygones, de polyèdres, et, en général, de figures à N dimensions, c'est la liaison algébrique entre les grandeurs non-statistiques de la superfigure, du "monde" qui joue le rôle primordial. C'est pourquoi on doit commencer la construction de la statistique de figures à N dimensions par la démonstration de cette formule liant les caractéristiques non-statistiques de la série des figures traitées statistiquement.

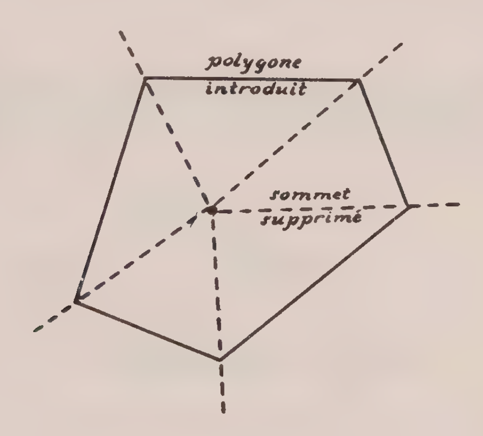


Fig. 4

Commençons par le cas des polygones. Remplaçons chaque sommet par un polygone auxiliaire d'après la fig. 4. Le nombre total f de polygones de notre série augmente naturellement; les grandeurs e et k augmentent aussi. Indiquons les grandeurs avant l'introduction des polygones auxiliaires par l'indice 1, les grandeurs après cette opération par l'indice 2.On a

$$\begin{aligned} e_1 &\longrightarrow e_2 = e_1\bar{\mu} = 2k_1 ; \\ f_1 &\longrightarrow f_2 = f_1 + e_1 ; \\ k_1 &\longrightarrow k_2 = k_1 + e_1\bar{\mu} = 3k_1 ; \end{aligned} \qquad (22)$$

Supposons qu'il existe une fonction φ liant e,f,k,; soit φ (e,f,k) = constante. Il est évident qu'après la transformation (22) la valeur de cette fonction ne changera pas

$$\varphi(e_1, f_1, k_1) = \varphi\left(2k_1, e_1 + f_1, 3k_1\right) \qquad (23)$$

C'est une équation fonctionnelle. Formellement on peut la résoudre en la différenciant. Mais il y a un danger, car ce n'est possible qu'en introduisant des hypothèses supplémentaires. Cependant, d'après M. G. Darmois(5) le problème peut être facilement résolu si on introduit les différences finies, par exemple

$$e_3 = e_1 + \Delta, \qquad f_3 = f_1 - \Delta \tag{24}$$

Ainsi on trouve facilement que

$$\varphi(e_1, f_1, k_1) = \varphi(e_1 + f_1, k_1) \tag{25}$$

Pour le détail sur l'application de cette méthode (voir G. Darmois). En appliquant encore une fois le même procédé on obtient:

$$\varphi(e_1, f_1, k_1) = \varphi(e_1 + f_1 - k_1) = \text{const.}, \quad e_1 + f_1 - k_1 = \text{const.}, \tag{26}$$

relation qui formellement coïncide avec la formule d'Euler et qui a été déjà appliquée dans la partie I.

§ 5 - Le même procédé est facilement applicable pour le cas de polyèdres. Introduisons autour de chaque sommet un polyèdre auxiliaire, (comme on le voit sur la fig.5). Comme précédemment,

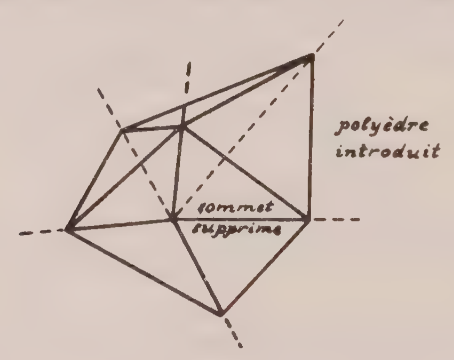


Fig. 5

appliquons l'indice 1 aux valeurs des grandeurs non-statistiques (valeurs caractéristiques du "monde") avant l'opération introduisant les polyèdres auxiliaires et l'indice 2 aux grandeurs après cette opération. Alors les valeurs avec l'indice 1 seront transformées dans les valeurs ayant l'indice 2 d'après le schéma

$$\begin{aligned} E_1 &\longrightarrow E_2 = E_1 \bar{\mu}_1 . \\ K_1 &\longrightarrow K_2 = K_1 + E_1 \bar{\mu}_2 \\ F_1 &\longrightarrow F_2 = F_1 + E_1 \bar{\mu}_3 \\ M_1 &\longrightarrow M_2 = M_1 + E_1 \end{aligned} \tag{27}$$

Supposant que la statistique de polygones a été déjà établie, la première ligne de (27) s'écrira aussi sous la forme

$$E_1 \longrightarrow E_2 = E_1 \bar{\mu}_2 - 2E_1 - E \bar{\mu}_3 \tag{28}$$

Une opération arbitraire ne pouvant changer une relation fondamentale, on a

$$\varphi(E_1, M_1, F_1, K_1) = \varphi(E_1\bar{\mu}_2 + 2E_1 - E\bar{\mu}_3, M_1 + E_1, F_1 + E\bar{\mu}_3, K_1 + E\bar{\mu}_2). \quad (29)$$

Le fait que le deuxième membre ne peut pas dépendre de μ_2 nous amène à la nécessité d'introduire dans la fonction la différence $E_1 - K_1$; le fait de l'indépendance de μ_3 introduit la somme $E_1 + F_1$.

Il en résulte $\varphi(E_1, M_1, F_1, K_1) = \varphi(2E_1 + F_1 - K_1, M_1 + E_1)$. (30)

En appliquant le procédé des différences finies 5) (voir les paragraphes précédents) on trouve définitivement:

$$\varphi(E_1, M_1, F_1, K_1) = \varphi(E_1 + F_1 - K_1 - M_1) - \text{const.},$$

et: (31)

$$E_1 + F_1 - K_1 - M_1 = \text{const.}$$

formule appliquée déjà plus haut pour établir la statistique de polyèdres. Sur la détermination de constantes dans (26) et (31), voir II-§6-D).

Pour exprimer plus systématiquement les résultats obtenus, introduisons les symboles symétriques. Soient M_i - les nombres des éléments à i dimensions; l'expression "éléments" étant employée dans le sens le plus général: M_0 - nombre de sommets, M_1-nombre d'arêtes etc.., enfin - pour le cas de figures à N dimensions, - M_{N+1} - nombre de "mondes" ($M_{N+1} = 1$).
Dans ces symboles on a:

$$\begin{array}{lll} \text{pour } N = 0, & M_0 = M_1 + 1 ; & M_1 = 1 ; \\ \text{pour } N = 1, & M_0 + M_2 = M_1 + 1 ; & M_2 = 1 ; \\ \text{pour } N = 2, & M_0 + M_2 = M_1 + M_3 + 1 ; & M_3 = 1 ; \\ \text{pour } N = 3, & M_0 + M_2 + M_4 = M_1 + M_3 + 1 ; & M_4 = 1 ; \end{array} \quad (32)$$

Deux premiers résultats sont évidents: un point dans un "monde" linéaire a deux"angles" , deux "côtés" ; et dans une surface - monde les figures linéaires forment une "série" ayant autant de sommets qu'il y a de figures. Les deux derniers de (32) représentent les lois d'Euler, classiques et pour le cas de quatre dimensions.

Nous venons de donner leur démonstration statistique.

§ 6 - Généralisons maintenant les résultats (32). Remarquons d'abord qu'on peut les représenter sous la forme des sommes

$$\begin{array}{lll} N = 0 ; \sum_{i=0}^{\infty} (-1)^i M_i = 1 ; & M_1 = 1 ; & M_2 = M_3 \ldots M_\infty = 0 \\ N = 1 ; \sum_{i=0}^{\infty} (-1)^i M_i = 1 ; & M_2 = 1 ; & M_3 = M_4 \ldots M_\infty = 0 \\ N = 2 ; \sum_{i=0}^{\infty} (-1)^i M_i = 1 ; & M_3 = 1 ; & M_4 = M_5 \ldots M_\infty = 0 \\ N = 3 ; \sum_{i=0}^{\infty} (-1)^i M_i = 1 ; & M_4 = 1 ; & M_5 = M_6 \ldots M_\infty = 0 \end{array} \quad (33)$$

Alors, on peut supposer que pour le cas général, il a:

$$\sum_{i=0}^{\infty}(-1)^i \underset{N}{M_i} = 1 \; ; \qquad \underset{N}{M_{N+1}} = 1 \; ; \quad \underset{N}{M_{N+2}} = \underset{N}{M_{N+3}} \dots \underset{N}{M_{\infty}} = 0 \qquad (34)$$

Ici nous avons introduit un deuxième indice auprès de M_i, à savoir $\underset{N}{M_i}$, indiquant par N le nombre de dimensions des figures considérées. La plus simple démonstration statistique de la formule(34) est celle que l'on peut faire à l'aide de la méthode de l'induction mathématique. Etant donné (33), supposons que la formule (34) est vraie pour N. Démontrons qu'elle sera vraie aussi pour N + 1. La démonstration suit les étapes ci-après.

A - Appliquons le principe mentionné de l'équivalence géométrique entre les propriétés d'une figure et d'un sommet. Dans ce qui suit, nous préférons ne conserver l'expression "sommet" que pour l'usage habituel; nous employons l'expression "étoile" pour indiquer l'ensemble des propriétés caractéristiques d'un sommet par analogie avec le concept "figure". De l'équation (33) qui est vraie comme nous l'avons supposé il découle que pour une "étoile" les équations

$$\sum_{i=1}^{\infty}(-1)^{i+1} \underset{N}{\bar{\mu}_i} = 1 \; ; \; \underset{N}{\bar{\mu}_{N+2}} = 1 \; ; \; \underset{N}{\bar{\mu}_{N+3}} = \underset{N}{\bar{\mu}_{N+4}} = \dots\dots \underset{N}{\bar{\mu}_{\infty}} = 0 \qquad (35)$$

sont vraies aussi. Par exemple, pour N = 2 on a:

$$\begin{aligned} &\underset{2}{M_0} - \underset{2}{M_1} + \underset{2}{M_2} - \underset{2}{M_3} = 1 \; ; \; \underset{2}{M_3} = 1 \; ; \; \underset{2}{M_4} = \underset{2}{M_5} = \dots\dots \underset{2}{M_{\infty}} = 0 \; ; \\ &\underset{2}{\bar{\mu}_1} - \underset{2}{\bar{\mu}_2} + \underset{2}{\bar{\mu}_3} - \underset{2}{\bar{\mu}_4} = 1 \; ; \; \underset{2}{\bar{\mu}_4} = 1 \; ; \; \underset{2}{\bar{\mu}_5} = \underset{2}{\bar{\mu}_6} = \dots\dots \underset{2}{\bar{\mu}_{\infty}} = 0 \, . \end{aligned} \qquad (36)$$

B - Considérons un système quelconque(une série de figures) M_i et remplaçons-le par le système M_i', en introduisant dans chaque sommet (point) une figure auxiliaire. On a alors:

$$\begin{aligned} &\underset{N+1}{M_0} \longrightarrow \underset{N+1}{M_0'} = \underset{N+1}{M_0}\,\underset{N}{\bar{\mu}_1} \; ; \\ &\underset{N+1}{M_1} \longrightarrow \underset{N+1}{M_1'} = \underset{N+1}{M_1} + \underset{N+1}{M_0}\,\underset{N}{\bar{\mu}_2} \; ; \\ &\text{-------------------------------------} \\ &\underset{N+1}{M_N} \longrightarrow \underset{N+1}{M_N'} = \underset{N+1}{M_N} + \underset{N+1}{M_0}\,\underset{N}{\bar{\mu}_{N+1}} \; ; \\ &\underset{N+1}{M_{N+1}} \longrightarrow \underset{N+1}{M_{N+1}'} = \underset{N+1}{M_{N+1}} + \underset{N+1}{M_0} \end{aligned} \qquad (37)$$

d'où naturellement suit l'invariance de la somme

$$\sum_{i=0}^{\infty}(-1)^i \underset{N+1}{M_i'} = \sum_{i=0}^{\infty}(-1)^i \underset{N+1}{M_i} \qquad (38)$$

D'après le procédé déjà appliqué, introduisons une fonction φ exprimant une relation invariante entre M_i. D'après (37) et (35) cette invariance sera exprimée par

$$\varphi\left(\frac{M_0}{N+1}, \frac{M_1}{N+1}, \dots \frac{M_{N+1}}{N+1}\right) = \varphi\left(\frac{M_0}{N+1}\frac{\bar{\mu}_1}{N}, \frac{M_1}{N+1} + \frac{M_0}{N+1}\frac{\bar{\mu}_2}{N}, \dots \frac{M_N}{N+1} + \frac{M_0}{N+1}\frac{\bar{\mu}_{N+1}}{N}, \frac{M_{N+1}}{N+1} + \frac{M_0}{N+1}(-1)^{N+1}\sum(-1)^i\frac{\mu_i}{N}\right) \quad (39)$$

C – Dans la formule (39) le deuxième membre ne dépend pas de μ. Utilisant ce fait et introduisant les différences finies[5] (voir plus haut), on trouve facilement que la fonction φ à la forme

$$\varphi\left(\frac{M_0}{N+1}, \frac{M_1}{N+1} \dots \frac{M_{N+1}}{N+1}\right) = \varphi\left(\sum_{i=0}^{\infty}(-1)^i\frac{M_i}{N+1}\right) \quad (40)$$

c'est-à-dire que

$$f\left(\sum_{i=0}^{\infty}(-1)^i\frac{M_i}{N+1}\right) = \text{const. et } \sum_{i=0}^{\infty}(-1)^i\frac{M_i}{N+1} = \text{const.} \quad (41)$$

Il ne reste qu'à déterminer la valeur de la constante dans cette formule.

D – Pour calculer cette valeur on peut prendre en considération les séries de figures les plus simples, parce que cette constante ne dépend pas, d'après la démonstration donnée, du choix de la série. C'est pourquoi considérons les "figures minimales" (les "simplex reguliers de W.I. Stringham[6]), figures ayant pour l'espace donné, le minimum de sommets. Pour l'espace à N dimensions ce minimum est évidemment égal à N + 1: pour déterminer un polygone sur une surface il est nécessaire d'avoir trois points ou plus; pour un polyèdre dans l'espace à trois dimensions, il est nécessaire d'avoir quatre sommets ou plus, etc. Ainsi pour un triangle on a:

$$M_0 = 3\,;\ M_1 = 3\,;\ M_2 = 1;\ \sum M_i(-1)^i = 1 \quad (42)$$

pour un tétraèdre:

$$M_0 = 4\,;\ M_1 = 6\,;\ M_2 = 4;\ M_3 = 1\,;\ \sum M_i(-1)^i = 1 \quad (43)$$

pour un "pentastère" („5-cell" de W.J. Stringham[6]) :

$$M_0 = 5\,;\ M_1 = 10\,;\ M_2 = 10\,;\ M_3 = 5\,;\ M_4 = 1;\ \sum M_i(-1)^i = 1 \quad (44)$$

Les formes correspondantes sont représentées dans la figure 6 où elles sont décomposées pour être représentables dans l'espace à N – 1 dimensions.

En général on a: nombre de lignes (arêtes) égal au nombre des combinaisons de (N+1) par deux $\frac{N\,(N+1)}{2} = C^2_{N+1}$,
nombre de surfaces (faces) égal au nombre des combinaisons de (N+1) par trois $= \frac{(N+1)\ N\ (N-1)}{6} = C^3_{N+1}$,

nombre de corps (polyèdres) égal au nombre des combinaisons de (N+1) par quatre = C^4_{N+1},
etc.., jusqu'au nombre de i-formes égal au nombre des combinaisons de (N+1) par i+1 = C^{i+1}_{N+1}

triangle:

tetraèdre:

pentastère:

Fig. 6

Alors on a en général $\underset{N}{M}_i = C^{i+1}_{N+1}$. Composant les sommes on trouve:

$$\sum_{i=0}^{N} (-1)^i \underset{N}{M}_i = 1-(1-x)^{N+1}_{x=1} = 1 , \qquad (45)$$

c'est-à-dire la valeur de la constante cherchée est déterminée: elle est égale à 1. Ainsi la formule (34) est établie.

Cette "méthode combinatoire"[ou cette "détermination combinatoire"]a été appliquée pour la première fois par B.Kirkmann[3], mais seulement dans le cas de trois dimensions, - voir aussi W.I. Stringham déjà cité.

La construction de toutes les autres formules de la statistique de figures à N dimensions revient à la simple énumération et à la construction des moyennes dans le sens des formules (3) et (15). Ainsi la construction abstraite de la statistique des figures étant en principe achevée, considérons qualques exemples d'application.

III. APPLICATIONS DE LA STATISTIQUE DE POLYGONES ET DE POLYÈDRES

§ 1 - La statistique de polygones et de polyèdres peut être appliquée à des problèmes très différents. Pour donner une idée sur le genre de problèmes où cette statistique est applicable, mentionnons d'abord un résultat purement négatif.

On peut s'attendre à ce que, non seulement les valeurs moyennes des nombres caractéristiques soient liés entre eux, mais qu'il en soit de même des distributions des valeurs de ces nombres parmi les figures d'une série. Cependant une dépendance entre les distributions n'existe pas. Pour démontrer cette affirmation purement négative, il suffit de donner des exemples des séries de figures ayant la même distribution d'une valeur caractéristique et différentes distributions des autres. Pour simplifier, donnons un exemple des séries de figures régulières, et notamment de polygones, pour faciliter les calculs et les représentations graphiques. Deux de telles séries sont représentées sur les figures 7 et 8. Au bas des mêmes figures on voit les distributions de n et de μ : Les distributions de μ sont les mêmes, tandis que celles de n sont différentes. Etant donné que pour le cas de polygones, on n'a que deux valeurs caractéristiques, nous nous persuadons qu'il n'existe aucune liaison univoque entre les distributions. Alors, pour le cas de polygones, on n'a que les formules

$$R\bar{n} + R\bar{\mu} = R2 + Rk \text{ , et } \overline{Rn} + \overline{R\mu} = R2 + Rk \qquad (46)$$

déjà citées et des limites

$$3 < \bar{n} < 6 \text{ ; } 3 < \bar{\mu} < 6 \text{ ; } 3 < n < \infty \text{ ; } 3 < \mu < \infty \qquad (47)$$

pour la variation de la valeur caractéristique.

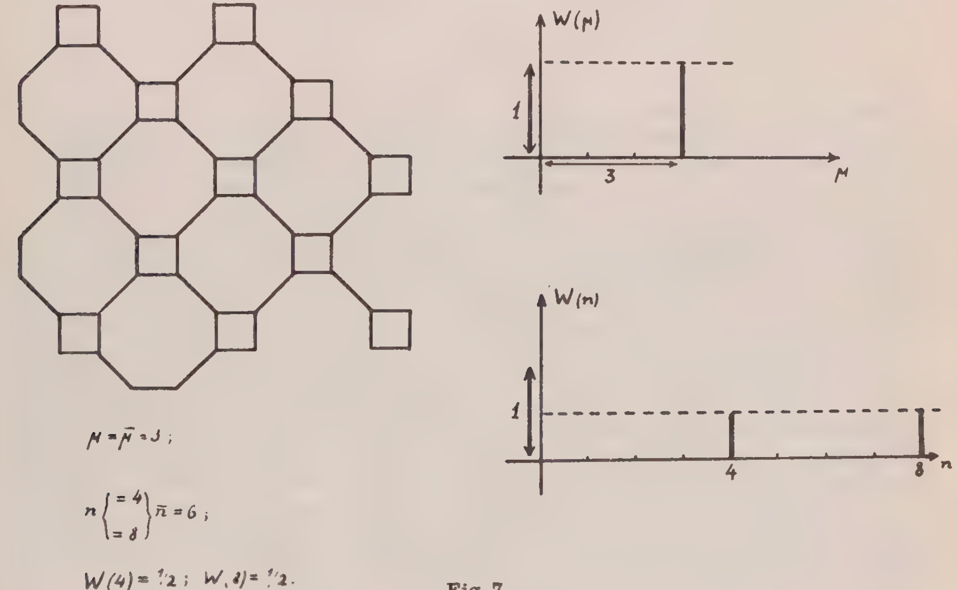


Fig. 7

On obtient facilement le même résultat pour les cas de plusieurs dimensions.

Alors on ne peut chercher que des applications où notamment les valeurs moyennes sont importantes. On peut classer tous ces problèmes d'application parmi deux types.

Dans les applications du premier type, la théorie statistique développée détermine les limites des domaines où on peut exiger les explications physiques, mécaniques,ou autres et ceux,

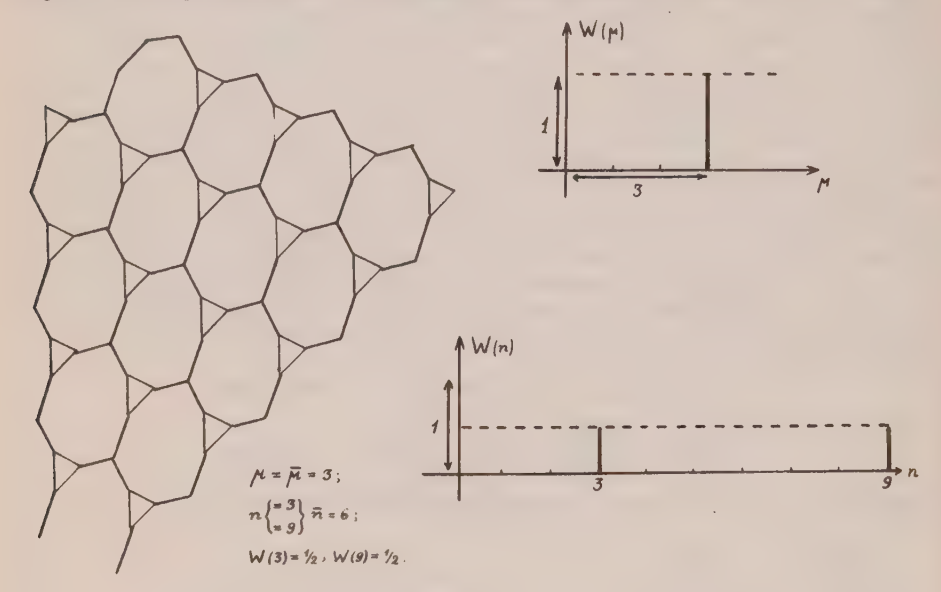


Fig. 8

au contraire, où nous n'avons rien que les faits purement mathématiques. Le problème de ce genre est, par exemple, le problème de sols polygonaux (voir (8)). On a cherché plusieurs fois à expliquer le fait de la prédominance de figures à six côtés parmi les polygones, formant le réseau des fissures sur un tel sol par plusieurs hypothèses: thermique, mécanique et autres. Cependant, comme nous le montrons dans la note citée, c'est complètement inutile. Du fait que deux fissures ne peuvent s'entrecroiser il découle que $\mu = 3$. De cela, à son tour, découle que $\bar{n} = 6$ et tout est, si l'on veut, expliqué. Demander dans ce cas une explication physique, c'est, pas plus et pas moins, demander une explication physique du fait que dans un triangle la somme des angles est égale à 180°. Plusieurs applications de la statistique de polygones et de polyèdres, aux problèmes de la géomorphologie (prismes basaltiques dans les laves (8), (9), etc....), astrophysique (9) (granulés sur la surface du soleil, etc..), météorologie, (9), (10)(nuage, etc...) et plusieurs autres sont encore du même genre. Dans la théorie des tourbillons (10),(11), les mêmes considérations simples peuvent être appliquées; cependant ici des applications plus profondes peuvent aussi avoir lieu. Nous les examinerons dans un article spécial.

§ 2 - Une autre classe d'applications de la statistique développée est représentée par les applications, où on cherche à déterminer une série ou, au moins, ces caractéristiques en partant des expériences sur une partie de cette série (3) . Ici on peut de

nouveau distinguer deux possibilités: 1° déterminer les caractéristiques statistiques d'un espace à N dimensions, d'après la section, c'est-à-dire d'après les données sur l'espace à (N-1) dimensions; 2° déterminer les caractéristiques d'une série (espace) par les mesures dans une partie limitée dans cet espace (déterminer toutes les caractéristiques d'une série ne connaissant que les caractéristiques statistiques). Le premier problème est le problème de base de la pétrographie mathématique, où on doit déterminer les propriétés d'une série des polyèdres (des cristaux et des espaces entre les cristaux) à partir des mesures sur une plaque mince, mesure des polygones, non des polyèdres.

Un problème analogue peut être imaginé pour la distribution de polygones.

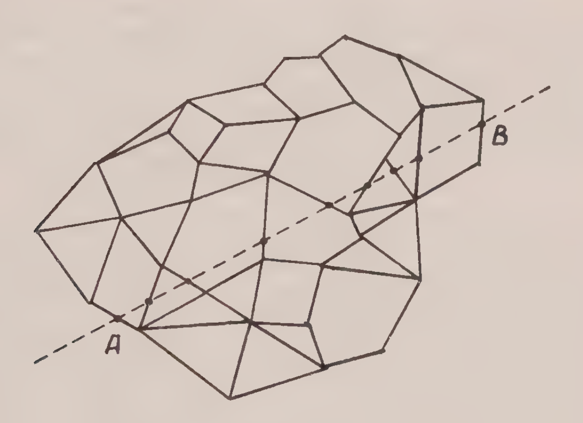


Fig. 9

Soit Δ une distribution de polygones dans un plan euclidien (voir fig.9). Cette série de polygones est caractérisée par $\overline{n}$ et $\overline{\mu}$. Supposons que nous ne connaissions rien sur cette distribution. Le problème est donc de déterminer Δ ou, au moins, $\overline{n}$ et $\overline{\mu}$ qui ne nous sont pas connues d'avance. De plus, nos mesures sont bornées à ce qu'on peut mesurer sur un segment AB. A l'aide de mesures dans cette "section" nous ne connaissons que m , le nombre des intersections des côtés avec ces segments .Comme dans toutes les mesures dans une section, on suppose que la section est représentative. La densité des intersections t, qui est égale à $t = m/AB$, peut moyennant que ce segment soit représentatif, déterminer le nombre de côtés des polygones se trouvant sur l'unité de surface; cette dernière densité est égale à $\alpha\ t^2$, où α est une constante statistique. C'est tout ce qu'on peut obtenir sans hypothèses quantitatives. Ce ne sont pas seulement les séries de polygones où nous nous trouvons dans l'impossibilité de déterminer la série sans des hypothèses supplémentaires; c'est également vrai pour toutes les séries de figures à nombre pair de dimensions.

Les hypothèses quantitatives, qu'on peut introduire ici, sont très nombreuses. Un exemple simple peut être donné en supposant que le nombre de polygones sur l'unité de surface est

connu; soit σ. Il est évident que le nombre de côtés de polygones sur l'unité de surface peut être exprimé de deux manières: en partant du nombre de polygones, il est égal à $\frac{\bar{n}\sigma^2}{2}$, en partant des mesures dans la "section" il est $\alpha\, t^2$. En comparant ces deux expressions et en appliquant la formule de base, on a:

$$\bar{n} = 2\alpha\left(\frac{t}{\sigma}\right)^2 ; \quad \bar{\mu} = \frac{2\alpha(t/\sigma)^2}{\alpha(1/\sigma)^2 - 1} \qquad (48)$$

§ 3 - Contrairement aux cas de séries de figures à nombre pair de dimensions, ce problème de détermination d'une série en partant de la section peut être complètement résolu[3] sans hypothèses quantitatives pour les cas de séries de figures au nombre impair de dimensions. Cependant les hypothèses "qualitatives" sont nécessaires, pas les hypothèses sur le nombre ("quantités") de choses, mais celles sur leurs "relations". Ceci pourrait être, en premier lieu, une hypothèse "angulaire" comme la relation trigonométrique de notre note [3], - voir ci-dessous.

Un représentant de séries de figures au nombre impair de dimensions est la série de polyèdres, problème cardinal de la pétrographie mathématique. Avant d'aborder ce problème, soulignons qu'il ne s'agit ici que de l'absence des hypothèses quantitatives - (on ne suppose rien sur la connaissance de telles ou telles autres valeurs numériques). Mais nous ne pouvons pas nous dispenser d'hypothèses qualitatives; et, outre la relation trigonométrique, des hypothèses que la section et les mesures sur elle sont <u>représentatives</u>. Tout ce qui suit est basé sur cette "représentativité" des mesures. On procède en évaluant d'abord la valeur $\bar{\mu}_{img}$, $\bar{\mu}$ imaginaire des polygones fictifs formés sur le plan de notre "section" par les intersections de côtés des polyèdres avec le plan de mesures. Evidemment $\bar{\lambda} = \bar{\mu}_{img}$. Alors une des caractéristiques est déterminée. Pour les connaître toutes, on doit encore tirer des mesures éventuelles la valeur d'une des autres caractéristiques. On peut par exemple déterminer à l'aide de mesures - ℓ^2 (densité des côtés de polygones fictifs dont nous venons de parler) et s^2 (densité de ces polygones mêmes). En analogie complète avec l'introduction de la constante statistique α (voir le paragraphe précédent), on peut introduire les constantes statistiques β_1 et β_2 qui nous permettent d'imaginer $\beta_1\ell^2$ (densité de faces de polyèdres) et $\beta_2 S^2$ (densité de polyèdres mêmes). Etant donné que la première grandeur peut être aussi représentée par $\frac{\bar{f}\beta_2 s^2}{2}$, on a:

$$\frac{\bar{f}\beta_2 s^2}{2} = \beta_1 \ell^2 , \text{ et } \bar{f} = 2\frac{\beta_1}{\beta_2}\left(\frac{\ell}{s}\right)^2 , \qquad (49)$$

ce qui nous détermine $\bar{f}$. Les grandeurs $\bar{\lambda}$ et $\bar{f}$ connues, on calcule les autres caractéristiques à l'aide des équations données plus haut et

$$\cos\left(\frac{\pi}{\bar{\mu}}\right) = \sin\left(\frac{\pi}{\bar{n}}\right) \cdot \sin\left(\frac{\pi}{\bar{\lambda}}\right) \qquad (50)$$

ce qui exprime "l'hypothèse qualitative" mentionnée au commencement du paragraphe. La relation (50) n'est exacte que pour les polyèdres réguliers. Elle peut être acceptée pour les polyèdres de tous types considérés en moyenne. En introduisant on commet une erreur d'ailleurs pas très grande si l'on ne considère pas les polyèdres très aplatis ou très allongés. Nous nous occupons de l'estimation de cette erreur, ainsi que d'autres formes éventuelles d'hypothèses quantitatives en autre endroit. Ici nous n'avons indiqué que le schéma général ; cependant ce calcul donne une idée des applications de la statistique développée aux problèmes de "sections".

§ 4 – Il nous reste encore à faire un court exposé sur les applications de la statistique de figures aux problèmes de la détermination d'une série (espace) en partant des mesures sur une de ses parties, partie ayant cependant le même nombre de dimensions que la série même [12] . Ici, il existe de même comme on l'a déjà souligné deux fois ci-dessus à l'occasion d'autres problèmes, une différence essentielle entre les séries de figures a nombres pair et impairs de dimensions.

Commençons par le cas des polygones. Considérons une série très grande de polygones; supposons qu'une partie de cette série est <u>représentative</u>, et mesurons (ou, pour mieux dire, comptons), $\bar{n}$ et $\bar{\mu}$ dans cette partie. L'hypothèse de la "représentativité" équivaut à admettre que les valeurs de $\bar{n}, \bar{\mu}$ de cette partie sont les mêmes que celles de $\bar{n}, \bar{\mu}$ pour toute la série. Ainsi en appliquant la formule (5), on calcule k , nombre de côtés dans toute la série caractéristique de la série entière. Connaissant k , on obtient facilement e, f, etc.. On voit que les mesures sur la partie représentative nous donnent des indications sur le "volume" du "monde". Pour les polygones sur un plan euclidien, la valeur de l'expression: $\frac{1}{\bar{n}} + \frac{1}{\bar{\mu}}$ est égale à $\frac{1}{2}$; c'est un "monde" infini et plan. La valeur numérique $\frac{1}{2}$ joue ici le même rôle que la valeur numérique π dans la géométrie usuelle. Si la valeur $\frac{1}{\bar{n}} + \frac{1}{\bar{\mu}}$ diffère de $\frac{1}{2}$, ceci indique que le "monde" n'est pas plan, en analogie avec le cas où le rapport $\frac{\text{circonférence}}{2\ (\text{rayon})}$ diffère de π. Au contraire, si on suppose que le plan est vraiment "plan", les mesures de $\bar{n}, \bar{\mu}$ nous amène à calculer la valeur de 2, en analogie complète avec le procédé par lequel on calcule π en jetant l'aiguille sur un cercle dessiné. Alors les mesures de $\bar{n}$, $\bar{\mu}$ sont équivalentes dans un certain sens aux mesures de la courbure.

En opposition avec tout cela, on ne peut rien obtenir à l'aide de telles mesures dans une partie de la série de polyèdres, même si cette partie est <u>représentative</u>. Ici, comme il s'ensuit de (20), ce ne sont que les valeurs relatives de E, F, K, M qui entrent dans les formules déterminant les grandeurs statistiques.

Deux hypothèses peuvent être énoncées:

Soit I : que la courbure de la série ("monde") de figures à nombre impair de dimensions n'est pas mesurable à partir des grandeurs statistiques.

Nous reviendrons à cette possibilité dans le §6.

Soit II : que, à la courbure donnée ne correspond aucune valeur déterminée du volume du "monde".

§5 - Cette dernière hypothèse s'accorde très bien avec le fait évident que dans les espaces pairs on peut imaginer l'infinité des figures régulières, tandis que les espaces impairs n'ont qu'un nombre fini de telles figures.

Imaginons par exemple que nous sommes sur la surface d'un polyèdre et que nous tentons de déterminer ce polyèdre à partir de mesures locales. Supposons pour fixer l'idée que les mesures dans un domaine représentatif nous ont donné: $\bar{n}=3$, $\bar{\mu}=3$. Alors, composant l'expression déterminant la courbure: $1/\bar{n} + 1/\bar{\mu} - 1/2$, on obtient: $1/\bar{\mu} + 1/\bar{n} - 1/2 = 1/6$. C'est-à-dire que notre polyèdre à 6 arêtes et par conséquent 4 faces: c'est un tétraèdre, sans aucun doute possible.

Au contraire, imaginons maintenant que nous sommes "sur" ou "dans" le volume d'un polystère (polytope de 4 dimensions). Nous mesurons:

a) toutes les grandeurs déterminant statistiquement une cellule, un élément: $\bar{n}$, $\bar{\mu}$, $\bar{e}$, $\bar{k}$, $\bar{f}$ (qui équivalent à la grandeur $\bar{n}$ déterminant statistiquement la cellule, l'élément dans le cas d'un polyèdre); et,

b) toutes les grandeurs déterminant statistiquement le point de la conjonction des cellules, des éléments: $\bar{\lambda}$, $\bar{\mu}$, $\bar{\mu}_1$, $\bar{\mu}_2$, $\bar{\mu}_3$ (qui équivalent à la grandeur $\bar{\mu}$ déterminant statistiquement le point de la conjonction des cellules des éléments dans le cas d'un polyèdre). Supposons que nous avons trouvé: $\bar{n}=3$, $\bar{\mu}=3$, $\bar{e}=4$, $\bar{k}=6$, $\bar{f}=4$, $\bar{\lambda}=3$, $\bar{\mu}_1=4$, $\bar{\mu}_2=6$, $\bar{\mu}_3=4$. Cependant, à partir de formules énumératives données dans II §2 et II §3, il est impossible de calculer M, E, K, F; on n'obtient que: $E/M = 1$; $K/M = F/M = 2$. On sait que "5-cell" de Stringham(6) a exactement les valeurs des rapports indiqués entre E, M, F, K. On peut par conséquent supposer, sans introduire de contradictions, que E=M=5, K=F=10 et que nous sommes "dans" ou "sur" un pentastère -„5-cell". Mais avec les mêmes raisons on peut supposer que la figure (le polystère) a $E=M=5\omega$ et $K=F=10\omega$ où ω est un nombre entier (ou même non-entier) arbitraire.

Tout cela est vrai si on ne dispose que des relations énumératives mentionnées. Le seul procédé à partir duquel on peut déterminer univoquement les nombres caractéristiques d'un polystère, est l'introduction plus ou moins arbitraire d'une relation du type (50), basée sur les raisons de "l'évidence géométrique". En opérant avec cette "évidence" (qui, en cas de 4 (et plus) dimensions est plus qu'arbitraire) on peut éliminer l'infinité de solutions des équations (20) et (31) homogènes en M, E, F, K, et parvenir dans chaque cas concret à une solution unique. Ainsi,on trouve,pour le cas de 4 dimensions, 6 polystères de Stringham.

Mais, nous soulignons ce fait, cette détermination univoque, dans les cas de 2ν dimensions, ne découle pas des relations énumératives; tandis qu'en cas de $2\nu+1$ dimensions toutes les grandeurs sont déterminables à partir de ces relations énumératives ici développées.

On peut avoir des opinions diamétralement opposées sur notre droit d'introduire, ou de non-introduire, les généralisations des relations non énumératives; par exemple les relations "angulaires" du type de (50). Toutefois, un fait reste toujours le même: cette généralisation de (5o) (- ou d'autres relations "angulaires" non énumératives) est nécessaire dans les cas ν pair et pas nécessaire dans les cas ν impair.

§6 - Reprenons maintenant l'hypothèse I(voir la fin du §4). On peut trouver à cette hypothèse plusieurs formes, formes nombreuses et très différentes. La plus naturelle est la suivante. Elle se base sur les raisons qui sont indiquées dans le paragraphe précédent comme appartenant à "l'évidence géométrique".

Il est "évident" que,dans une série de polyèdres dans un espace à trois dimensions, espace euclidien ou courbé, il ne se produit pas de changements des valeurs moyennes des angles auprès des sommets et auprès des arêtes, si l'on conserve les valeurs des grandeurs caractéristiques: μ, n, λ, e, k, f, μ_1, μ_2 μ_3, et,si l'on ne change que les longueurs relatives des arêtes. Cependant une affirmation réciproque n'est pas valable. Pour établir l'impossibilité de cette affirmation réciproque il suffit d'un exemple

La série des octaèdres réguliers, où chaque arête appartient aux quatre polyèdres ($\lambda=4$), est formée par 24 octaèdres c'est un "24-cell" de Stringham[6] construit en partant de "l'évidence géométrique". Si chaque arête n'appartient pas aux quatre mais à un nombre plus petit de polyèdres ($\lambda=3$ ou λ entre 3 et 4), le nombre des octaèdres réguliers dans le "monde" serait encore plus petit: $\mathcal{M}<24$. Ceci pourrait être illustré par les tétraèdres qui, touchant les arêtes au nombre de trois ($\lambda=3$), forment un "monde" de 5 figures ("5-cell" de Stringham) et qui, touchant ces arêtes au nombre quatre ($\lambda=4$), forment un "monde" de 16 figures ("16-cell"). Alors le "monde" pour les octaèdres réguliers touchant entre eux au nombre moins de quatre ($\lambda<4$) contient moins de 24 figures.

Considérons maintenant les octaèdres non-réguliers dont chacun est formé par les parties des deux cubes joints par la face ABCD. Soient O et O' les centres de ces cubes. L'octaèdre considéré est OABCDO' . Il est évident que le "monde" formé par ces polyèdres, est infini ($M=\infty$). Pour les arêtes AB, BC, CD et DA le nombre λ est égal à 4; pour les autres arêtes il est égal à trois. On voit que la valeur moyenne de λ se trouve entre 3 et 4 ; $\lambda<4$.

Alors deux séries des polyèdres sont construites,- séries ayant les mêmes valeurs des paramètres caractéristiques: μ, n, λ, e, k, f, μ_1, μ_2, μ_3, - l'une,des polyèdres réguliers, l'autre des polyèdres non-réguliers (OABCDO'). Pour la première série

$M < 24$, tandis que pour la deuxième $M=\infty$; les grandeurs E, K, F, diffèrent pour ces deux séries dans la même proportion. Alors, dans le cas de l'espace au nombre impair de dimensions, la constance des grandeurs non-statistiques M, etc. détermine la conservation des grandeurs statistiques (voir ci-dessus) même si l'on change les longueurs des arêtes. Au contraire, ces grandeurs statistiques laissent libre le choix des grandeurs non-statistiques.

Rien de pareil n'existe dans les espaces au nombre pair de dimensions. Par exemple pour le cas d'un plan, le nombre f de polygones dans le "monde" est complètement déterminé par les nombres caractéristiques n et μ . D'aucune manière il ne dépend des longueurs de côtés des polygones formant la série. La même chose est vraie pour les séries des "5", "6",...."Q-cell" dans l'espace à quatre dimensions, etc...

Tout cela est étroitement lié avec la forme que prend la formule (34) dans les cas de N pair ou impair[12]. Pour N augmentant, cette formule prend alternativement la forme avec un terme constant ou sans ce terme; alternativement cette formule contient, soit les valeurs absolues de M_i, soit leurs valeurs relatives.

Ces résultats[12] peuvent être définitivement formulés sous la forme suivante: en partant de grandeurs statistiques, les grandeurs M sont déterminables dans les espaces à nombre pair de dimensions et indéterminables dans les espaces à nombre impair de dimensions. C'est pourquoi il convient de distinguer:

a) Les espaces statistiquement indéterminables, au sens géométrique les espaces "non-propres", "hypersurfaces"; et

b) Les espaces déterminables, au sens géométrique "propres".

BIBLIOGRAPHIE

1) Minkowski H. Théorie der konvexen Körper. III und IV Kap. Volumen und Oberfläche, §§ 4 - 8. Gesammelte Abhandlungen, 1911. Teubner. Berlin.

2) Steiner J. Über parallele Flächen. Akademische Monatsberichte, 1840, p.114-118, Berlin

3) Matschinski M. Statistique de polygones et de polyèdres. Comptes rendus Ac. Sci. t. 237(1953), p.1310 .

4) Schoute. Mehrdimensionale Geometrie. II Teil. Die Polytope. Sammlung Schubert, Leipzig 1905.

5) Darmois G. Bulletin of the International statistical Institute, Volume XXXIII, Part II, p.79.

6) Stringham W.I. Journ. of math. 3 (1880), p.1

7) Kirkmann B. Mancherster Mem. (2), 12, (1855), p. 70

8) Matschinski M. Sur les formations hexagonales. Bul. Soc. Geol.Fr., 6° Série, t.2, fasc.7-9.

9) Wasiutinski, J. Studies in hydrodynamics and structure of stars and planets.-Edit.J.Dybwad, Oslo 1946.

10) Matschinski M. Systèmes de tourbillons sujets à certaines conditions de minima. Application à la formation des rides éoliennes, des dunes, etc... - Annales de Géophysiques, n°4, p.402 (1952)

11) Bénard H. Les tourbillons cellulaires dans une nappe liquide. Revue générale des Sciences, vol.XI, pp.1261 et 1309 (1900)

12) Matschinski M. Sur la géométrie statistique.

Composition - Impression
par les Procédés
" DACTYL - SPÉCIAL "

ACHEVÉ D'IMPRIMER
LE 10 OCTOBRE 1954
SUR LES PRESSES DE
J. & R. SENNAC
54, Fbg Montmartre, 54
PARIS (9e)

N° d'Imprimeur : 5.668

J. & R. SENNAC, Imprimeurs
54, Fg Montmartre, PARIS (9e)